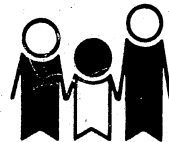


ADVANCES IN CONSUMER RESEARCH

VOLUME V

**PROCEEDINGS OF THE EIGHTH
ANNUAL CONFERENCE OF THE
ASSOCIATION FOR CONSUMER RESEARCH**



**Edited by
H. KEITH HUNT**

1978 copyright

© ASSOCIATION FOR CONSUMER RESEARCH

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.

International Standard Book Number: 0-915552-03-5

International Standard Serial Number: 0098-9258

Hunt, H. Keith, Editor. *Advances in Consumer Research*, Volume V. (Ann Arbor: Association for Consumer Research, 1978).

ORDER FORMS ARE PROVIDED ON PAGE 786

ASSOCIATION FOR CONSUMER RESEARCH - EIGHTH ANNUAL CONFERENCE

October 13 - 16, 1977 - Chicago, Illinois

CONFERENCE COMMITTEE

H. Keith Hunt, Conference Chairman
Brigham Young University

* * * *

Jay D. Lindquist
Western Michigan University

William D. Perreault, Jr.
University of North Carolina

Earl W. Morris
Iowa State University

William L. Wilkie
University of Florida

Robert W. Pratt, Jr.
Heublein, Inc.

John Eighmey, Arrangements
Northwestern University

* * * *

Fred D. Reynolds
University of Georgia

Anthony F. McGann
University of Wyoming

Mary Winter
Iowa State University

George E. Brosseau, Jr.
National Science Foundation

EXECUTIVE COMMITTEE

Harold H. Kassarian, President
University of California, Los Angeles

Jerome B. Kernan, President-Elect
University of Cincinnati

David M. Gardner, Past President
University of Illinois

Kenneth L. Bernhardt, Executive Secretary
Georgia State University

William B. Locander, Newsletter Editor
University of Houston

Richard J. Lutz, Treasurer
University of California, Los Angeles

H. Keith Hunt, Conference
Brigham Young University

John Eighmey, Conference
Northwestern University

George H. Haines, Jr., Membership
University of Toronto

Mary Jane Schlinger, Publication
University of Illinois at Chicago Circle

James R. Bettman, Director
University of California, Los Angeles

Jerry C. Olson, Director
Pennsylvania State University

Gerald Zaltman, Director
University of Pittsburgh

Robert W. Pratt, Jr., JCR Policy Board
Representative, Heublein, Inc.

ADVISORY COUNCIL

Gwen Byers
Cornell University

Theodore F. Dunn
Benton & Bowles Advertising

Robert Ferber
University of Illinois

Stephen A. Greyser
Harvard University & Marketing Science Institute

Edward Heiden
U.S. Consumer Product Safety Commission

Stanley D. Shores
Procter & Gamble

Jay Schmeideskamp
Gallup Organization, Inc.

James R. Taylor
University of Michigan

Joseph N. Uhl
Purdue University

Frank D. Walker
Walker Research, Inc.

LIAISON REPRESENTATIVE FROM NATIONAL SCIENCE FOUNDATION

George E. Brosseau, Jr.

PREFACE

ADVANCES IN CONSUMER RESEARCH, VOLUME V, contains the papers presented at the Eighth Annual Conference of the Association for Consumer Research. The ACR annual conference provides a forum for the presentation of papers and discussion of topics which advance the state of the art in consumer research. This Proceedings contains 133 papers representing the contributions of 231 authors.

ACR annual conferences traditionally include three types of sessions: (1) competitive paper sessions featuring papers selected through a competitive double-blind reviewing process in which members of the Association serve as reviewers, (2) special topic sessions in which the topic and proposed participants of the session are competitively judged by the conference program committee, and (3) workshop sessions in which the topic and general area of discussion are competitively judged by the conference program committee. The papers appear in the proceedings grouped by the type of session of which they were part.

Conferences which make a contribution to the state of the art require the contribution of many people. First must go thanks to the authors and to the organizers of the special topic and workshop sessions. The lasting value of the conference and of the Proceedings will be due to their individual efforts. The reviewers also play a special role, both the program committee who review the workshop and special topics proposals and the competitive paper reviewers. Both groups are listed elsewhere in the Proceedings. Also deserving thanks are the many secretaries who provided camera-ready copy of papers for the Proceedings. And I especially appreciate the helpful hints of previous editors of these proceedings who shared their experience with me and encouraged me in this work.

As anyone who has ever edited a Proceedings knows, behind the editor is a super secretary who somehow keeps everything in order, cajoles and begs threatens to get papers in on time, and retypes and retypes and retypes. My special thanks go to Willyne Forsyth who was the person behind the conference details and behind the editing of this Proceedings.

Finally, I appreciate the support and understanding of my wife, Carolyn, during the weekends, late night, and even all night sessions while the conference and Proceedings were being put together. Her handling of family matters gave me the freedom to be involved in the highly enjoyable experience of being program chairman and Proceedings editor.

H. Keith Hunt

Editor

Graduate School of Management and
College of Business

Brigham Young University

Provo, Utah 84602

TABLE OF CONTENTS

Association For Consumer Research Officers and Conference Committee	iii
Preface	iv
H. Keith Hunt, Brigham Young University	
Presidential Address	xiii
Harold H. Kassirjian, University of California, Los Angeles	

PART I: COMPETITIVE PAPERS

"On the Social Psychology of Giving: Door-in-the-Face and When Even a Penny Helps"	1
Peter H. Reingen, University of South Carolina	
"Selling a City: An Experimental Study of the Communication Effects of Message Tone"	5
Richard F. Yalch, University of Washington Margaret C. Dempsey, University of Washington	
"The Relative Effectiveness of One-Sided and Two-Sided Communication for Mass Transit Advertising"	12
Linda Golden, University of Texas, Austin Mark I. Alpert, University of Texas, Austin	
"Cola Preferences: Disguised Taste vs. Brand Evaluations"	19
Thomas J. Stanley, Georgia State University	
"Nonmetric Unidimensional Scaling of Consumer Preferences for Proposed Product Designs"	22
Murphy A. Sewell, State University of New York, Albany	
"Understanding University Choice: A Multi-Attribute Approach"	26
Ronald Vaughn, Bradley University Joseph Pitlik, Bradley University Behram Hansotia, Bradley University	
"An Exploratory Study of the Influences of Usage Situation on Perceptions of Product-Markets"	32
Rajendra K. Srivastava, University of Pittsburgh Allan D. Shocker, University of Pittsburgh George S. Day, University of Toronto	
"Assessing the Effects of Visible Consumption on Impression Formation"	39
Russell W. Belk, University of Illinois, Urbana-Champaign	
"Evoked Set Size as a Function of Number of Choice Criteria and Information Variability"	48
Joseph A. Belonax, Jr., Bowling Green State University Robert A. Mittlestaedt, University of Nebraska, Lincoln	
"Consumer-Oriented Versus Advertiser-Oriented Language: Comprehensibility and Salience of the Advertising Message"	52
Martin R. Lautman, Associates for Research in Behavior, Inc. Larry Percy, Gardner Advertising Company	
"Who Is the Deal Prone Consumer?"	57
Robert Blattberg, University of Chicago Thomas Buesing, University of Chicago Peter Peacock, Wake Forest University Subrata Sen, University of Rochester	
"Comparative Advertising: Issues and Problems"	63
Michael Etgar, State University of New York, Buffalo Stephen A. Goodwin, State University of New York, Buffalo	
"Mediating Effects of Cognitive Responses to Advertising on Cognitive Structure"	72
Jerry C. Olson, Pennsylvania State University Daniel R. Toy, University of Washington Philip A. Dover, Dartmouth College	
"The Organization of Product Information in Memory Identified by Recall Times"	79
Eric J. Johnson, Carnegie-Mellon University J. Edward Russo, University of Chicago	

"Low Involvement Versus High Involvement Cognitive Structures"	87
John L. Lastovicka, Temple University	
David M. Gardner, University of Illinois, Urbana-Champaign	
"A Study of the Interface Between Attitude Structure and Information Acquisition Using a Questionnaire-Based Information-Display Sheet"	93
Morris B. Holbrook, Columbia University	
Karl A. Maier, Ketchum, MacLeod & Grove, Inc.	
"Alternative Approaches to Assessing the Quality of Self Report Data"	99
Robert A. Hansen, Ohio State University	
Carol A. Scott, Ohio State University	
"An Evaluation of Telephone Sampling Designs"	103
E. Laird Landon, Jr., University of Houston	
Sharon K. Banks, University of Oregon	
"Whether to Agree-Disagree or Disagree-Agree: The Effects of Anchor Order on Item Response"	109
David Sheluga, Purdue University	
Jacob Jacoby, Purdue University	
Brenda Major, Purdue University	
"Nearest Neighbor Analysis: Inferring Behavioral Processes from Spatial Patterns"	114
Sanford L. Grossbart, University of Nebraska, Lincoln	
Robert A. Mittelsteadt, University of Nebraska, Lincoln	
Gene W. Murdock, University of South Dakota	
"Durable Accumulation: An Examination of Priority Patterns"	119
Robert F. Lusch, University of Oklahoma	
Edward F. Stafford, Jr., University of Oklahoma	
Jack J. Kasulis, University of Oklahoma	
"An Investigation Into the Causal Links Between Attribution Schema and Decision-Making"	126
Richard Mizerski, University of Cincinnati	
Stephen Green, University of Cincinnati	
"Some Issues in Designing Consumer Information Studies in Public Policy".	131
R. Bruce Hutton, University of Denver	
Dennis L. McNeill, Federal Trade Commission	
William L. Wilkie, University of Florida	
"Potential Contributions of Consumer Research to Antitrust Decision Making"	138
Paul N. Bloom, University of Maryland	
"Expectancy-Discrepant Information: Does Format Influence Effects?"	145
Debra L. Scammon, University of Utah	
"Shopping Time and Leisure Time: Some Preliminary Cross-Cultural Comparisons of Time-Budget Expenditures"	151
Douglass K. Hawes, University of Wyoming	
Sigmund Grønmo, Norwegian Fund for Market and Distribution Research	
Johan Arndt, Norwegian School of Economics and Business Administration	
"Personal Values: A Cross Cultural Assessment of Self Values and Values Attributed to a Distant Cultural Stereotype"	160
J. Michael Munson, University of Santa Clara	
Shelby H. McIntyre, University of Santa Clara	
"Correlates of Information Gathering and Evoked Set Size for New Automobile Purchasers in Norway and the U.S."	167
R. Neil Maddox, University of Missouri, St. Louis	
Kjell Grønhaug, Norwegian School of Economics and Business Administration	
Richard E. Homans, University of Missouri, St. Louis	
Frederick E. May, University of Missouri, St. Louis	
"Selecting Pareto Optimal Subsets from Multiattribute Alternatives"	171
James B. Wiley, Wayne State University and University of Florida	
"Four Methodological Problems in Multi-Attribute Attitude Models"	175
Michel Laroche, Simmons College	
"Measured Attribute Weights Can Make A Difference"	180
Richard A. Werbel, University of Illinois, Chicago Circle	
"Assessing the Impact of Increased Product Safety on Consumer Utility"	186
Stephen C. Burnett, Indiana University	

"Transportation Attitudes Over Time: A Longitudinal Approach"	194
Mark I. Alpert, University of Texas, Austin Linda L. Golden, University of Texas, Austin	
"Energy Conservation, Price Increases and Payback Periods"	201
William H. Cunningham, University of Texas, Austin Brondel Joseph, University of Illinois, Chicago Circle	
"An Experimental Examination of Deception in Labeling: Consumer Research and Public Policymaking"	206
Philip G. Kuehl, University of Maryland Robert F. Dyer, George Washington University	
"Evaluation of Food Labeling Policies Through Measurement of Consumer Utility"	213
Roger Best, University of Arizona Jim McCullough, University of Arizona	
"Consumer Processing and Use of Supplementary Drug Label Information"	220
F. Robert Dwyer, University of Arizona	
"A Multiple Discriminant Analysis Approach to the Development of Retail Store Positioning"	227
Lawrence J. Ring, University of Virginia Charles W. King, Purdue University	
"A Hierarchical Model of Source Effect in Retail Newspaper Advertising"	235
James G. Barnes, Memorial University of Newfoundland	
"An Examination of Consumer Grocery Store Choice: Considering the Attraction of Size and the Friction of Travel Time"	243
Daniel J. Brown, Oregon State University	
"Perceptions of Unfair Marketing Practices: Consumerism Implications"	247
Gerald Zaltman, University of Pittsburgh Rajendra K. Srivastava, University of Pittsburgh Rohit Deshpande, University of Pittsburgh	
"A Comprehensive Study of Consumer Satisfaction with Durable Products"	254
Stephen B. Ash, Indiana University	
"Consumer Response to Dissatisfaction with Services and Intangibles"	263
Ralph L. Day, Indiana University Muzaffer Bodur, George Mason University	
"Catastrophe Theory as a Model for Describing Consumer Behavior"	273
Terence A. Oliva, Louisiana State University Alvin C. Burns, Louisiana State University	
"Testing Competing Models of Consumer Decision Making in the Preventive Health Care Marketplace"	277
Richard L. Oliver, University of Iowa Philip K. Berger, University of Kentucky	
"An Examination of an Alternative Form of the Behavioral Intention Model's Normative Component"	283
Michael J. Ryan, Columbia University	
"The Market for Personal Growth Services"	290
Philip Kotler, Northwestern University Lenore Borzak, Northwestern University	
"Sparks and Tucker Revisited: A Reanalysis and Replication"	295
Charles M. Schaninger, University of Massachusetts V. Parker Lessig, University of Kansas	
"Perception of Marital Roles in Decision Processes: Replication and Extension"	300
E. H. Bonfield, Temple University	
"An Examination of Innovative Communicators, Opinion Leaders and Innovators for Men's Fashion Apparel"	308
Elizabeth C. Hirschman, University of Pittsburgh William O. Adcock, Georgia State University	
"Individual Time Orientation and Consumer Life Style"	315
Robert B. Settle, San Diego State University Pamela L. Alreck, San Diego State University John W. Glasheen, San Diego State University	
"Belief Systems and the Differential Role of the Self-Concept"	320
George E. Belch, University of California, Los Angeles	

"Masculinity and Feminity Factors in Product Perception and Self Image"	326
James W. Gentry, Kansas State University	
Mildred Doering, Kansas State University	
Terrence V. O'Brien, Kansas State University	

PART II: SPECIAL TOPICS PAPERS

"Attitude Maturation: Changes in Related Belief Structures over Time"	333
Jerry C. Olson, Pennsylvania State University	
Philip A. Dover, Dartmouth College	
"Carryover Effects of Corrective Advertising"	343
Alan G. Sawyer, Ohio State University	
Richard J. Semenik, University of Utah	
"Attitudes as Actions: Social Identity Theory and Consumer Research"	352
Barry R. Schlenker, University of Florida	
"A Functional Approach to Consumer Attitude Research"	360
Richard J. Lutz, University of California, Los Angeles	
"New Directions in Attitude Research: A Critical Evaluation"	370
Joel B. Cohen, University of Florida	
"Short and Long Term Cost Consequences of Government Regulations to Consumers"	377
Rachel Dardis, University of Maryland	
"Industry Response to Government Regulations"	380
David Hillier, American Textile Manufacturers Institute	
"Issues in Analyzing Consumer Satisfaction/Dissatisfaction with Clothing and Textiles"	383
George B. Sproles, University of Houston	
Loren V. Geistfeld, Purdue University	
"The Process of Setting Standards and Some Current Issues in Standards"	392
L. J. Sharman, Consumer Product Safety Commission	
"Human Values: A Historical and Interdisciplinary Analysis"	396
C. Joseph Clawson, University of Southern California	
Donald E. Vinson, University of Southern California	
"Values and Consumption Patterns: A Closed Loop"	403
James M. Carman, University of California, Berkeley	
"The Europeanizing of America: A Study of Changing Values"	408
Donald L. Kanter, University of Southern California	
"Consumers' Reactions to Sex in TV Commercials"	411
Deborah K. Johnson, Needham, Harper & Steers Advertising, Inc.	
Kay Satow, Needham, Harper & Steers Advertising, Inc.	
"Sex on Television, More or Less"	415
William S. Rubens, National Broadcasting Company	
"Men and Women on the Tube: Sex and Sex-Stereotypes on Television"	419
Susan Hesselbart, Florida State University	
"Energy Conservation and Travel Behavior"	422
Jeffrey S. Milstein, U. S. Department of Energy	
"Consumer Oriented Transportation Planning: An Integrated Methodology for Modeling Consumer Perceptions, Preference and Behavior"	426
Alice M. Tybout, Northwestern University	
John R. Hauser, Northwestern University	
Frank S. Koppleman, Northwestern University	
"Functional Measurement Analysis of Spatial and Travel Behavior"	435
Jordan J. Louviere, University of Wyoming	
Irwin P. Levin, University of Iowa	
"Extended Conjoint Analysis with Intensity Measures and Computer Assisted Interviews: Applications to Telecommunication and Travel"	440
John R. Hauser, Northwestern University	
Steven M. Shugan, University of Rochester	

"Consumer Research Implications of Random Coefficient Models"	448
N. Scott Cardell, Charles River Associates	
Ricardo Dobson, Charles River Associates	
Frederick C. Dunbar, Charles River Associates	
"Attitude-Behavior Changes in a Before-After Mode Choice Situation"	456
Mary Lynn Tischer, Federal Highway Administration	
Carl G. Shea, Federal Highway Administration	
"Halo Effects in Marketing Research: Review and Prognosis"	465
Neil E. Beckwith, University of Pennsylvania	
Harold H. Kassarian, University of California	
Donald R. Lehman, Columbia University	
"A Measure of Halo"	468
Joel Huber, Columbia University	
William James, University of Alabama	
"Halo Effects and Location Preferences"	474
William L. James, University of Alabama	
Forrest S. Carter, Georgia Institute of Technology	
"Econometric Estimation of Halo Effect: Single vs. Simultaneous Equation Models"	477
Albert C. Bemmar, Purdue University	
Joel C. Huber, Columbia University	
"A Study of the Amount of Halo in the Perceptions of Automobiles"	481
William L. Moore, Columbia University	
William L. James, University of Alabama	
"Empirical Evidence of Halo Effects in Store Image Research by Estimating True Locations"	485
Neil E. Beckwith, University of Pennsylvania	
Victor Kubilius, University of Pennsylvania	
"Identifying and Defining Consumer Needs Using Human Factors and Market Research Techniques"	494
Thomas Cannon, Design Factors	
Ronald W. Hasty, Colorado State University	
"Infant, Child and Teenager Anthropometry for Product Safety Design"	499
Richard G. Snyder, University of Michigan	
Lawrence W. Schneider, University of Michigan	
Clyde L. Owings, University of Michigan	
"Safety Research: You and Your Lawn Mower"	508
Ann M. RameySmith, National Bureau of Standards	
Val J. Pezoldt, National Bureau of Standards	
Jay J. Persensky, National Bureau of Standards	
"Children's Conceptions of Medicine: The Role of Advertising"	515
Thomas S. Robertson, University of Pennsylvania	
John R. Rossiter, University of Pennsylvania	
Terry C. Gleason, University of Pennsylvania	
"Consumer Behavior and Nutrition: Preventive Health Perspectives"	518
M. Venkatesan, University of Oregon	
"Environmental Protection for the Non-Smoker: Consumer Behavior Aspects of Encouraging Non-Smoking"	521
Lawrence H. Wortzel, Boston University	
Roberta Clarke, Boston University	
"Consumer Behavior and Health Care Change: The Role of Mass Media"	525
Thomas S. Robertson, University of Pennsylvania	
Lawrence H. Wortzel, Boston University	
"A Review of Formal Theories of Consumer Socialization"	528
Deborah L. Roeder, Northwestern University	
Nicholas M. Didow, Northwestern University	
Bobby J. Calder, Northwestern University	
"Children's Consumer Information Processing: Representation of Information from Television"	535
Ellen Wartella, Ohio State University	
Daniel B. Wackman, University of Minnesota	
Scott Ward, Harvard University and Marketing Science Institute	
"The Effects of TV Messages for High and Low Nutritional Foods on Children's Snack and Breakfast Food Choices"	540
Marvin E. Goldberg, McGill University	
Gerald J. Gorn, McGill University	
Wendy Gibson, Stanford University	

"Psychometric Characteristics of Behavioral Process Data: Preliminary Findings on Validity and Reliability"	546
Jacob Jacoby, Purdue University	
Robert W. Chesnut, Columbia University	
Wayne D. Hoyer, Purdue University	
David A. Sheluga, Purdue University	
Michael J. Donahue, Purdue University	
"Subjects' Information Processing in Information Display Board Studies"	555
David C. Arch, University of California, Los Angeles	
James R. Bettman, University of California, Los Angeles	
Pradeep Kakkar, University of Pennsylvania	
"Eye Fixations Can Save the World: A Critical Evaluation and a Comparison Between Eye Fixations and Other Information Processing Methodologies"	561
J. Edward Russo, University of Chicago	
"Verbal Protocols and Direct Observation of Supermarket Shopping Behavior: Some Findings and a Discussion of Methods"	571
John W. Payne, Duke University	
E. K. Easton Ragsdale, University of Chicago	
"The Use of the Active Profile Evaluation Paradigm in Studying Judgment Processes"	578
Peter Wright, Stanford University	
Peter Rip, Stanford University	
"Chronometric Analysis: An Introduction and an Application to Low Involvement Perception of Advertisements"	581
Meryl P. Gardner, Carnegie-Mellon University	
Andrew A. Mitchell, Carnegie-Mellon University	
J. Edward Russo, University of Chicago	
"Motivation-Need Theories and Consumer Behavior"	590
W. Fred van Raaij, Tilburg University	
Kassaye Wandwossen, Tilburg University	
"Information About Consumer Problems by the Policy-Delphi Procedure: A New Method for Evaluating the Marketing System"	596
Solveig Wikstrom, Stockholm University	
"The Perception of Foreign Products in France"	603
Gary Baumgartner, University of Grenoble	
Alain Jolibert, University of Grenoble	
"Exposure to Cultural Activities and Opinion Leadership"	606
Eric Langeard, University of California, Berkeley	
M. Crousillat, Universite d'Aix-Marseille	
R. Weisz, Universite d'Aix-Marseille	
"Imagery-Eliciting Strategies: Review and Implications of Research"	611
Kathy Lutz, University of Southern California	
Richard J. Lutz, University of California, Los Angeles	
"Visual Imaging Ability as a Mediator of Advertising Response"	621
John R. Rossiter, University of Pennsylvania	
Larry Percy, Gardner Advertising Company	
"Cognitive Response, Imagery, and Scripts: What is the Cognitive Basis of Attitude?"	630
Bobby J. Calder, Northwestern University	
"Participation in Organizational Buying: Some Conceptual and Methodological Problems"	635
Kjell Grønhaug, Norwegian School of Economics and Business Administration	
"Research Approaches to Multi-Participant Decision Processes"	639
David T. Wilson, Pennsylvania State University	
"Observations of Buyer and Seller Transactions"	643
Arch G. Woodside, University of South Carolina	
James L. Taylor, Texas A & M University	
"Narrowing the Conceptual-Empirical Gap in Organizational Buying Behavior"	653
Robert E. Spekman, University of Maryland	
Bobby J. Calder, Northwestern University	
"On the Interface Between Organizational and Consumer Buying Behavior"	657
Yoram Wind, University of Pennsylvania	

"A Comparative Analysis of Determinant Attributes in Retail Store Selection"	663
Stephen J. Arnold, Queen's University Sylvia Ma, University of Toronto Douglas J. Tigert, University of Toronto	
"A Pragmatic Approach for Retail Fashion Monitoring"	668
Lawrence J. Ring, University of Virginia	
"Fashion Preferences and Store Patronage: A Longitudinal Study"	675
George B. Sproles, University of Houston	
"Segmenting Local Markets for Entertainment Services: The Case of Discotheques"	682
Daniel A. Emenheiser, Oklahoma State University George B. Sproles, University of Houston	
"Local Retail Market Monitoring: Strategic Implications"	688
Charles King, Purdue University	
"Cross-Cultural Research Methodology as a Case of Construct Validity"	693
W. Fred van Raaij, Tilburg University	
"Research for Consumer Policy Formulation: An Exploration of Dimensions of Household Economic Management"	702
Johan Arndt, University of Missouri, St. Louis Oddvar Holmer, Norwegian School of Economics and Business Administration	
"Session Summary: Cross-Cultural Life Style Research"	704
M. Venkatesan, University of Oregon	
"Comments on Cross-Cultural Consumer Research"	705
Johan Arndt, University of Missouri, St. Louis	
"Inferential Belief Formation in the Cue Utilization Process"	706
Jerry C. Olson, Pennsylvania State University	
"Self-Perception Processes in Consumer Behavior: Interpreting One's Own Experiences"	714
Carol A. Scott, Ohio State University	
"A Two-Stage Theory of Information Processing in Persuasion: An Integrative View of Cognitive Response and Self-Perception Theory"	721
Alice M. Tybout, Northwestern University Brian Sternthal, Northwestern University Bobby J. Calder, Northwestern University	
"Cue Utilization in Product Perception"	724
Robert E. Burnkrant, University of California, Berkeley	

PART III: WORKSHOP PAPERS

"Fad Food Use Among Anglo- and Mexican-Americans: An Example of Research in Consumer Behavior and Home Economics"	730
Joel Saegert, University of Texas, San Antonio Eleanor Young, University of Texas, San Antonio Merry Mayne Saegert, University of Texas, San Antonio	
"The Problems, Pitfalls, and Opportunities in Interdisciplinary Applied Consumer Research"	734
J. Barry Mason, University of Alabama Helen Goetz, University of Alabama	
"Source Similarity and Fashion Newness as Determinants of Consumer Innovation"	738
Karen Kaigler-Evans, University of Texas, Austin Clark Leavitt, Ohio State University Lois Dickey, Ohio State University	
"Clothing Decisions: A Decision Process Analysis of Focused Group Interviews"	743
Roger D. Blackwell, Ohio State University Jo Ann Schickel Hilliker, Ohio State University	
"Research Priorities"	750
Elizabeth Y. Davis, U. S. Department of Agriculture R. J. Ritchie, Virginia Polytechnic Institute and State University Francille M. Firebaugh, Ohio State University	

"Administration and Funding of Interdisciplinary Consumer Research"
Keith N. McFarland, University of Minnesota

"Experiences in Interdisciplinary Consumer Research"
Sarah L. Manning, Purdue University

"Interdisciplinary Research Needs and Opportunities in Consumer Behavior and Education" (Group Report)
J. Barry Mason, University of Alabama
A. William Gustafson, University of Alabama
Mary Harrison, University of Florida
Norleen Ackerman, University of Wisconsin
Roger Swagler, University of Tennessee
Rose Davis, University of Kentucky
Beatrice Petrich, University of Wisconsin
Nancy Miller, University of Wisconsin
E. Thomas Garman, Virginia Polytechnic Institute and State University
Robert Mittlestaedt, University of Nebraska, Lincoln
Terrje Assum, Cornell University

"Priority of Textile and Clothing Problems for Interdisciplinary Consumer Research" (Group Report)
Gloria Williams, University of Minnesota
Betty Crown, University of Alberta
Marjorie Wall, University of Guelph
Billie G. Murphy, University of North Carolina, Greensboro
Lois Korslund, University of Missouri, Columbia

"Interdisciplinary Research Goals for Food, Human Nutrition, and Health" (Group Report)
Jim McCullough, University of Arizona
Karen Morgan, Michigan State University
Joel Saegert, University of Texas, San Antonio

"Energy Management for Citizens in a Contemporary World" (Group Report)
Merlene Lyman, South Dakota State University
Maryann Paynter, University of Kentucky
Emma Auer, Purdue University
Raedene Combs, University of Nebraska, Lincoln
Patricia Tripple, University of Nevada, Reno
Judy Van Name, University of Delaware

Conference Program
Authors' Addresses
Reviewers of Competitive Papers
Order Form

PRESIDENTIAL ADDRESS, 1977:
ANTHROPOMORPHISM AND PARSIMONY

Harold H. Kassirjian, University of California, Los Angeles

This is the eighth meeting of the Association for Consumer Research and the seventh time that members of the association have had to suffer through a farewell address by a president. On this particular occasion, I am particularly fortunate for in the audience sit previous presidents of this association and therefore at least they appreciate that if such addresses are insufferable to the listener, they are even more threatening to the speaker. I am here because Keith Hunt ordered me to give a presidential address at this place and at this time with the simple request that it be significant, visionary, important, memorable, and interesting: a milestone in the consumer research literature. What he meant was, "keep it mercifully short." Not too sure of what a visionary milestone was, I turned to previous addresses.

In 1970 our first leader, and father of this association, Jim Engle, stood here and spoke about the developing organization and its exciting future. Bob Perloff continued in the same vein. In 1973 Joel Cohen was optimistic and excited perhaps because the Journal of Consumer Research had been conceived during his administration and would soon be born. Cohen, however, did comment on his perceptions of the trivial nature of the problems we had been studying.

Bob Pratt again sat back and offered us his perspective of ACR, where it had been and the challenge ahead for the fledgling adolescent. That year will best be remembered for the wave after wave of laughter caused by a film he presented. By the time Jack Jacoby presented his sharp perspective and criticism we had matured, we had our own literature, we had our own research tradition, we had several media to disseminate our work.

And so, the tradition of a presidential address that was to be somewhat different from those of other associations began to emerge. As Gardner pointed out last year, rather than the presentation of a major research inquiry, rather than a significant contribution to the research literature that is common in our sister association; our presidential address was to be an overview, a perspective, a self criticism of our field and our association.

And now it is my turn. I would like to focus on two problems I see in the field of consumer behavior: that of anthropomorphism and our lack of concern about parsimony.

Anthropomorphism

First let me turn to anthropomorphism. As you know, pet owners often see in their animals a thinking, reasoning, information processing creature--a cat that ambles with disdain, an arrogant dog, a wise goldfish. More than one psychologist experimenting in animal laboratories has wondered whether his highly trained rat is thinking out a response, pondering a decision or has become "lost in thought." And children often wonder about the emotional feelings of a flower or the memories of a rock that has seen so many centuries pass before his (or is it her?) senses. To project human activities, emotions, motivations, and adjectives onto animals or inanimate objects is an error we all make, even in our professional work.

I can't help but wonder, if much too often we do not project too much onto the consumer: because we care about brands, or packaging, or information processing, we assume that the average consumer also cares. As consumer researchers, this group happens to be highly educated, extremely intelligent, and as individuals we pride ourselves on being wise, rational consumers not burdensomely influenced by advertising campaigns, or skillfully deceptive labels. If anyone is a "thinking, reasoning, problem solving organism capable of transforming, storing, and evaluating sensory inputs", this audience must qualify as that type of person. For if humans can be classified into motivational or personality types, this audience must qualify as the cognitive man.

And from that, we have assumed, perhaps much too often, that other people, consumers in general, are also extremely wise, problem solving rational organisms every time the purchase decision is made for a pound of coffee, a tube of tooth paste, or a can of peas. We have often assumed that labels are read, products are compared, prices are utilized, and a rational decision is made, when we know it is not usually so.

Of course, there is no question that under some conditions for the purchase of some goods and services, the consumer is sometimes a thinking, information processing individual. Important or expensive, high risk, high involvement, psychologically or socially or ego related products, may under certain conditions, lead to highly sophisticated actions and decisions. Thus our complex theories, and our model of the cognitive man occasionally seems to be rewarding.

But, of the dozens, if not hundreds, of mundane decisions made each day by the average consumer, I wonder just how many are important to him, just how many are significant or high involvement decisions. For example, the work by Tom Robertson, and Mike Ray, and of course, the classic studies by Krugman, and several papers at this conference clearly indicate that behavior toward the low involving product is quite different than behavior toward the high involvement, high risk, important product. And the world is mostly full of insignificant decisions and unimportant solutions.

Yet, because to us a brand of toilet paper may be important, we have projected that belief onto the consumer. A one percent share of market can mean thousands of dollars of net profit to a company. That one brand has greater shearing strength or is more quickly biodegradable, or has more sheets per roll may be quite important to the consumerist, to the environmentalist, or to the readers of Consumer Reports. That panel data on the purchase of toilet paper over a seventeen year period exists may be titillatingly exciting to the academic researcher, but the average consumer who blithely purchases, consumes, and discards the product, most likely could care less. Unconcernedly he or she makes the purchase, switches brands, ignores commercials, and worries about the important decisions in his life and not the purchase of toilet paper.

For these countless thousands of insignificant decisions that are made by the consumer, to assume a thinking, reasoned, attitudinally influenced decision may well be a classic example of anthropomorphism.

Parsimony

In a similar vein, we have developed complex theories in the attempt to understand, explain, and predict the behavior of the consumer, as had psychology many years before. In psychology, the all encompassing extremely complex views of Sigmund Freud were called upon as well as those of Adler, Jung, Fenichel, and Horney; and, of course, Clark Hull and Tolman and Kurt Lewin.

But, in time, psychology turned from the complex theories to the much simpler middle range theories that Merton had been calling for. Lately, in psychology, an even simpler theory has been quite dominating. From clinical psychology to counselling, from social psychology to learning, we see more and more the impact of behaviorism and the work of Skinner, stressing not the mystical beliefs of the ancient Greeks or of the turn of the 20th century Germans, but rather the simple concepts of repetition, reward, reinforcement, and shaping behavior.

To this point, the work in consumer behavior has not followed that pattern, from the all encompassing view to the middle range theory to the simpler theory -- but rather has been quite the opposite in its short history. For example, the original work by Howard in his marketing management textbook started with a very uncomplicated view of consumer behavior theory. Behavior could simply be explained by drive, response, cue, and reinforcement. In less than a decade, this simple parsimonious view has been expanded, developed, flow-charted, sophisticated, and complicated. Meanwhile, in marketing and advertising, from the simple behavioristic views that sales or advertising effects were related to repetition and reinforcement, we found researchers turning to middle range theories such as cognitive dissonance, and personality, perceived risk, multi-attribute attitude models, and attribution theory. By stringing several of these middle range theories together, the flow-chart models were to emerge.

And in our journey from the simple to the more and more complex, we lost our basic theoretical need for parsimony. Three years ago, Frank Bass, frustrated that the complex theories we created could not predict much of behavior, proclaimed in a Journal of Marketing Research article that, "it will never be possible to provide good predictions of individual choice behavior for separate occasions," and that man is substantially stochastic. Bass assumed that there is a homonculus in the brain that throws dice, and that behavior is dependent upon that stochastic turn of events. Perhaps Bass's view is too simplistic, and certainly it is not new, for it has been around since the early days of the mathematical learning theorists three decades ago. Nevertheless, the view deserves more of a hearing than it received in consumer behavior this time around. But apparently it was just too simple for us. Our cherished middle range theories, our complex deterministic views of reality were not prepared to accept a simple approach. It may be true that stochastic approaches have not predicted the behavior of the consumer in the marketplace very well, but they have had no less success than the unparsimonious approach of the grand theories of behavior. Having been personally schooled in psychoanalytic theory, having been weaned on Gestalt and Lewinian Field theory, and never much having liked the views of Bush and Estes, of Mosteller and Burke, I find it quite difficult to defend stochastic thinking and to condemn a deterministic view of reality. Certainly I do not believe that a Bass or Skinner can explain the selection of a spouse, the purchase of a home, the ownership of a new car or the decision to commit suicide; but it is possible that with the

unimportant, uninvolved, insignificant, minor decisions that are made in the marketplace everyday, we do not need a grand theory of behavior. In unimportant, uninvolved, unemotional, non-cognitive, low commitment instances, unparsimonious complex explanations are often just not necessary. Although it is quite hard to accept the extremely simplistic views of the mathematical theorists and stochastic models, perhaps we also do not need the complication of intervening variables and hypothetical constructs for explanatory power when studying low commitment purchases. Skinner's shaping behavior is perhaps more than enough in the low commitment marketplace; reinforcement and repetition may well be sufficient.

As Markin and Nayarana pointed out in a 1975 ACR paper, "the fact that so much consumer behavior is shaped in an operant fashion perhaps even refutes the notion of the cognitive consumer. Such a traditional concept may well reflect a tragic misunderstanding of our notions regarding consumer's intelligence, rationality, and intellectual autonomy."

Although the Markin position is quite extreme, if we divide the world into important, high involvement, high commitment decisions and into insignificant trivial decisions, it may well be that the theories to explain these two kinds of behavior are quite different. Whereas complex deterministic theories involving series of intervening constructs are necessary for the high involvement case; for the insignificant ones, those with relatively low risk, simple Skinnerian behaviorism, operant conditioning, or maybe even stochastic thinking may be more than sufficient.

It is too bad that so few of us have chosen to study learning theory in consumer behavior in recent years, and so few of us have turned to the study of low involvement behavior. Whether we like it or not, low involvement, low risk, the unimportant, is what much of consumer behavior is all about. Whether or not we anthropomorphize our own values onto the consumer, and whether or not we chose to be unparsimonious in our thinking, the facts do not change that under most conditions, for most types of goods and decisions, the behavior of the consumer is just not important from his point of view.

There is a subset of us that is, in fact, not really interested in consumer behavior, per se; but rather is interested in human behavior; and happened to choose human behavior in the marketplace as their specialty. It is here that the grand theories of consumer behavior properly reside. And, these people are making some important contributions to the world of science and knowledge.

But, for the rest of us, our strength and our contributions may be even greater. We, not pure economists, or pure psychologists, are the experts on the thousands of insignificant and trivial decisions in this world. It is here that we have made our greatest contributions to date and here lies our greatest potential for future contributions.

There is a lot more in the world that is low risk and low involvement than there is major and important. But these are decisions nonetheless. It is we that know more than any other sub-set in the scientific community on this topic. Perhaps if we have it, we should flaunt it.

ON THE SOCIAL PSYCHOLOGY OF GIVING: DOOR-IN-THE-FACE
AND WHEN EVEN A PENNY HELPS¹

Peter H. Reingen, University of South Carolina

Abstract

Four strategies of inducing people to comply with a request to donate money were investigated in a field experiment: A concession approach where a requester first asked a target person for an extreme favor and, after being refused, retreated to a smaller favor (the one that was desired from the outset); an approach that avoided the dilemma of small requests by legitimizing rather than directly requesting compliance with a small request; a combination of the previous two approaches; and an approach utilizing the standard plea for funds. The prediction that the standard approach would be the least effective was confirmed. However, contrary to predictions, the approaches involving a concession did not significantly increase compliance with a (third) request for subsequent help. Possible processes that could explain these results are discussed.

Introduction

There has been a recent upsurge of interest in methods for obtaining behavioral compliance with minimal pressure. The door-in-the-face technique, first formally investigated and confirmed by Cialdini, Vincent, Lewis, Catalan, Wheeler and Darby (1975), has received particular attention. This technique is based on the simple notion that the probability of obtaining compliance with a request that has a low a priori probability of agreement can be increased substantially by first inducing noncompliance with an extreme request. Cialdini et al. (1975) and Cialdini and Ascani (1976) argue that a requester's movement from an initial, extreme request to a second, more moderate one is viewed by the target person as a concession. Based on the societal rule of reciprocation of concessions that states "You should make concessions to those who make concessions to you," normative pressures are brought to bear that tend to compel a target person to move from his/her initial position of noncompliance with the extreme request to a position of compliance with the more moderate second request. Cialdini and his coworkers found this rejection-then-retreat approach effective in increasing compliance with the rather large request of chaperoning a group of juvenile delinquents on a day trip (Cialdini et al. 1975) and increasing compliance with a request to donate blood (Cialdini and Ascani, 1976).

One of the purposes of this study was to test the concession procedure in a request context other than the ones employed in previous research. Subjects (college students) were asked in a naturalistic setting to contribute money to the Heart Association in accordance with one of four conditions. In the first, the control condition, a standard request for funds was made. In the second, the extreme-then-donation request condition, subjects were first asked to perform an extreme favor (to give three dollars every month for at least one year) and, having refused, were asked the standard request. It was predicted that subjects exposed to the concession procedure would be significantly more likely to decide to contribute than those exposed to only the

standard request. This part of the study, then, simply replicated the work by Cialdini et al. In so doing, it was hoped to (a) test the generality of the concession procedure as a compliance induction mechanism, (b) obtain further evidence concerning the hypothesized mediator of the effect, (c) apply the technique in a context that is of considerable interest to both practitioners and consumer researchers, and (d) investigate the power of the technique by employing student subjects.

The purpose of the third condition was to test the relative effectiveness of the concession procedure by including in the experimental design a procedure that did not involve a concession. In this third condition, the standard request was followed by the sentence, "Even a penny will help." While the research evidence (e.g., Wagner and Wheeler, 1969) suggests a strong relationship between a subject's compliance with a request and the cost of performing the request, the clear dilemma of small requests per se is that they produce high levels of compliance but low levels of reward for a requester (Cialdini and Schroeder, 1976). The phrase "Even a penny will help," however, implied that a small contribution was acceptable without suggesting, as small requests normally do, that it was what the requester desired. In other words, by implying that a small favor is acceptable but not necessarily desirable, a requester would make it difficult for a target to refuse and would make it unlikely that a target would offer a small donation (Cialdini and Schroeder, 1976). In the fourth condition, the phrase "Even a penny will help" was added to the extreme-then-donation request manipulation. Since this approach would appear to reap the benefits of both the concession approach and the even-a-penny-will-help strategy but not the disadvantages of small requests, it was expected to be an especially effective compliance induction mechanism. That is, it was predicted that subjects exposed to the last approach would be significantly more likely to decide to contribute than those exposed to just the request addendum, "Even a penny will help," but that each of these two approaches would be superior to the standard request. In addition, based on the above, it was hypothesized that the average donation would not differ significantly between conditions.

Finally, the fund-raising context provided an opportunity to examine another kind of compliance of considerable interest, namely compliance with a request for subsequent help. Subjects in the extreme initial request conditions and a control condition were asked to become a fund-raising volunteer for the Heart Association. It was predicted that the subjects in the extreme initial request conditions would be more likely to comply with this (third) request than the control subjects. This expectation was based on evidence produced by studies of bargaining behavior suggesting that a concession strategy in negotiation leaves the target of such a strategy with more positive feelings about the interaction outcome (e.g., Benton, 1973), and on more direct evidence provided by Cialdini and Ascani (1976) showing that subjects in an extreme-then-blood donation request condition were more likely to agree to donate blood again than control subjects.

¹The author would like to thank the Heart Association of South Carolina and especially Mary Wright for her considerable assistance. This paper reports a portion of a larger study.

Method

Subjects

Subjects were 160 (80 male and 80 female) students at the University of South Carolina. Only those students were selected who were walking alone along university walkways during the hours 10 a.m. to 6 p.m., and no subjects known to an experimenter were selected.

Procedure

The naive experimenters, four male and four female college students, approached only same-sex subjects who met these conditions. Experimenters, thoroughly instructed in training sessions, were equipped with the identification badges, information brochures and donation envelopes in fund-raising drives by the Heart Association of South Carolina. An experimenter initiated interaction by introducing him- or herself as representing the Heart Association. After the common introductory remarks, subjects were randomly assigned to one of five conditions according to a treatment schedule that varied randomly across experimenters. An experimenter completed exactly four replications per condition.

Donation-request-only control condition. Subjects in this condition (n=32) were asked by the experimenter to comply with the request to donate money. Specifically, the experimenter said:

As part of our annual campus fund-raising drive, I'm collecting money for the Heart Association. Would you be willing to help by giving a single donation?

If a subject consented, the experimenter took the donation and put it in the donation envelope. Thus, two measures of helping were taken: (a) whether a subject gave a contribution, and (b) size of contribution.

Extreme-then-donation request condition. Subjects in this condition (n=32) were initially asked to perform a very large favor which all subjects refused. Specifically, the experimenter said:

As part of our annual campus fund-raising drive, we're currently asking students to become involved in our Long Term Donor Program. Long term donors are those students who pledge to give 3 dollars per month for a period of at least 12 months. Would you be willing to enroll in our Long Term Donor Program?

(After the subject declined, the experimenter continued.) Well, maybe you'd be interested in another activity we're asking students to participate in. I'm also collecting money.....

The remainder of the procedure was identical to that of the donation-request-only control.

Even-a-penny condition. Subjects in this condition (n=32) were delivered the standard (control) plea for funds, but the experimenter added, "Even a penny will help." The remainder of the procedure was identical to that of the donation-request-only control.

Extreme-then-donation request, even-a-penny condition. Subjects in this condition (n=32) were administered the procedure identical to that of the extreme-then-donation request condition, except that "Even a penny will help" was added to the standard plea. All subjects but one refused to comply with the extreme initial request.

Volunteer-request-only control condition. Subjects in this condition (n=32) were asked by the experimenter only to comply with the request to become a fund-raising

volunteer. Specifically, the experimenter said:

As part of our annual fund-raising drive, the Heart Association is looking for volunteers to solicit funds from the public. Would you be interested in becoming a Heart Association fund-raising volunteer?

If a subject consented, the experimenter wrote down his/her telephone number and indicated that the Heart Association would contact the subject if his/her help was needed. Thus, the dependent measure was the subject's verbal compliance. The subjects in the extreme initial request conditions were also approached with this volunteer request but after they had received the donation request.

Results

Donation Request

A preliminary Chi Square analysis on frequency of donation within each condition showed no significant differences due to sex of the target subject (and thus sex of the experimenter), with levels of significance ranging from .28 to 1.00, and no significant differences between experimenters, with levels of significance between .36 and .89.² Hence, the compliance data were relatively free of sex effects and experimenter effects, and the subsequent analysis could therefore be performed on collapsed data.

Table 1 presents the results for frequency of donation, total amount donated, and average amount donated. The predictions concerning compliance rates seemed most appropriately tested by a set of three orthogonal contrasts.

TABLE 1

Donation Results

Condition	% of Compliance	\$Total	\$Average
Donation-request-only	19(6/32)	3.07	.51
Extreme-then-donation	34(11/32)	6.75	.61
Even-a-penny	47(15/32)	4.64	.31
Extreme-then-donation, even-a-penny	44(14/32)	5.60	.40

The first tested the donation-request-only control against the combination of the extreme-then-donation request condition, the extreme-then-donation, even-a-penny condition, and the even-a-penny condition. It was predicted that the three experimental conditions would be superior to the control in compliance to donate money, and this was confirmed ($X^2 = 4.52$, d.f. = 1, $p = .03$). The second contrasted the extreme-then-donation, even-a-penny condition and the even-a-penny condition. The prediction that the extreme-then-donation, even-a-penny approach would be superior to the even-a-penny treatment was not confirmed ($X^2 = 0.00$). The last compared the extreme-then-donation and the combination of the extreme-then-donation, even-a-penny and the even-a-penny groups and also found no significant difference ($X^2 = .65$, d.f. = 1, $p = .42$).

With respect to mean donation per active contributor, no significant difference was expected. While this prediction was confirmed ($F = 2.08$, d.f. = 3/42, $p = .12$),

²All Chi Square tests reported in this paper were corrected.

there was a consistent tendency for the "even a penny will help" additions to produce a somewhat smaller average contribution. However, an a posteriori test based on Scheffé's procedure involving the relevant comparison resulted in a p of only .25.

Volunteer Request

Table 2 presents the proportion of subjects in the three relevant conditions who were willing to comply with the volunteer request. Contrary to expectations, it is clear that the principal orthogonal comparison (volunteer-request-only control versus the combination of the extreme-then-donation and extreme-then-donation, even-a-penny conditions) is highly insignificant.

TABLE 2
Volunteer Results

Condition	% of Compliance
Volunteer-request-only	6(2/32)
Extreme-then-donation	0(0/32)
Extreme-then-donation, even-a-penny	9(3/32)

Discussion

These results are of significant practical value for anyone wishing to elicit compliance with a request to donate money. The experimental conditions were significantly more effective in producing compliance with the donation request than the control. In addition, the average amount donated did not differ significantly among the conditions. Thus, it is not surprising that along the practical dimension of total funds obtained, the even-a-penny condition produced 1.5 times, the extreme-then-donation request, even-a-penny condition 1.8 times, and the extreme-then-donation request condition 2.2 times that of the donation-request-only control. That these results are consistent with the ones produced by Cialdini and his coworkers attests to the generality of the compliance induction mechanisms. However, the requests in this investigation as well as in the Cialdini et al. studies were altruistic ones. Thus, the effects have been shown to occur only under the somewhat limited set of conditions where the requests are prosocial in nature and are made in a no-delay situation. Evidence provided by Reingen and Kernan (1977b) suggests that the concession strategy may be ineffective in contexts where "selfish" rather than "socially desirable" requests are involved. A situation where selfish requests are involved would hardly resemble a bargaining interaction. Similarly, studies by Reingen and Kernan (1977a) and Snyder and Cunningham (1975) which employed time-delayed requests suggest that the concession approach may actually result in less compliance. These results are, of course, favorable to the concession viewpoint in that a requester's act of concession may not be as apparent when the requests are time-delayed.

Aside from these more practical aspects, the present findings have interesting conceptual implications as well. As far as the extreme initial request conditions are concerned, the findings provide support for a concession-based mediator of the compliance effects, but they as well as the ones obtained by Cialdini et al. in no way ultimately confirm the concession model. Other explanations may exist. For example, Cann, Sherman and Elkes (1975) suggest an alternative explanation based on dissonance theory. When a subject in an extreme initial

request condition refuses to comply with an initial request for help in a social cause, he probably feels some conflict about this which he can easily resolve by complying with the immediately-following second request. Another possibility is a perceptual contrast effect. By initially asking an extreme request, the second request may be perceived as smaller than it would have been if it alone had been presented. Interestingly enough, the major surprise of this study, which was the ineffectiveness of the extreme initial request conditions to produce a greater compliance with a request for subsequent help (i.e., volunteering as a fund raiser), can be easily explained by these two alternative interpretations but not by the concession model. From a dissonance viewpoint, the second request may have given ample opportunity for dissonance reduction. As far as the perceptual contrast explanation is concerned, the (third) request for subsequent help may have been perceived as large as the extreme (first) request, as is indicated by the very low compliance rate in the volunteer-request-only control group.

A second surprising aspect of the present data was the ineffectiveness of the extreme-then-donation request, even-a-penny condition to produce a greater compliance to donate money than the even-a-penny condition. Since the only difference between the two conditions was that the former involved a concession and the latter did not, no significant gain resulted from the concession approach. It appears that the addendum to the standard request, "Even a penny will help," is sufficient enough by itself to make it difficult for a target person to refuse to donate money.

References

- A.A. Benton, "Reactions to Demands to Win From an Opposite Sex Opponent," *Journal of Personality*, 41 (1973), 430-442.
- A. Cann, S.J. Sherman, and R. Elkes, "Effects of Initial Request Size and Timing of a Second Request on Compliance: The Foot in the Door and the Door in the Face," *Journal of Personality and Social Psychology*, 32 (1975), 774-782.
- R.B. Cialdini, J.E. Vincent, S.K. Lewis, J. Catalan, D. Wheeler, and B.L. Darby, "Reciprocal Concessions Procedure for Inducing Compliance: The Door-in-the-Face Technique," *Journal of Personality and Social Psychology*, 31 (1975), 206-215.
- R.B. Cialdini and K. Ascani, "Test of a Concession Procedure for Inducing Verbal, Behavioral, and Further Compliance With a Request to Donate Blood," *Journal of Applied Psychology*, 61 (1976), 295-300.
- R.B. Cialdini and D.A. Schroeder, "Increasing Compliance by Legitimizing Paltry Contributions: When Even a Penny Helps," *Journal of Personality and Social Psychology*, 34 (1976), 599-604.
- P.H. Reingen and J.B. Kernan, "Compliance With an Interview Request: A Foot-in-the-Door, Self-Perception Interpretation," *Journal of Marketing Research*, 14 (1977), in press (a).
- P.H. Reingen and J.B. Kernan, "Foot-in-the-Door or Door-in-the-Face? Further Evidence Toward a Resolution of the Paradigms," Center Paper #64, Center for Marketing Studies, University of South Carolina, (1977), (b).
- M. Snyder and M.R. Cunningham, "To Comply or not to Comply: Testing the Self-Perception Explanation of the Foot-in-the-Door Phenomenon," *Journal of Personality and Social Psychology*, 31 (1975), 64-67.

C. Wagner and L. Wheeler, "Model, Need, and Cost Effects in Helping Behavior," Journal of Personality and Social Psychology, 12 (1969), 111-116.

SELLING A CITY: AN EXPERIMENTAL STUDY OF THE
COMMUNICATION EFFECTS OF MESSAGE TONE

Richard F. Yalch, University of Washington
Margaret C. Dempsey (student), University of Washington

Abstract

Residents and nonresidents of Seattle were exposed to persuasive appeals varying in message tone but all urging them to live within the city limits. The results support the use of positively and neutrally worded messages, but suggest that negatively worded appeals be used only under limited circumstances.

Introduction

Despite widespread praise because of its many favorable attributes, Seattle, as many other large cities, experienced a constant decline in its population and lost businesses to its suburbs during the past few years. Government officials attributed this partially to a negative image of urban living presented in news media accounts of urban problems and they sponsored a series of television commercials extolling the virtue of living within the city limits. The objective was to present "a number of really positive things going on around town the people never hear about" (Seattle Turns to Net TV, 1977). Concurrently, a county transportation agency tried to increase the demand for public transportation with an advertising campaign criticizing individuals who commute by automobile. The campaign was sufficiently negative toward automobile drivers (the slogan was "I'm no dummy, I take the bus") to motivate a public protest from the Executive Vice President of the Automobile Club of Washington which included his recommendation that the agency switch to a "positive" appeal (Dummies Who Drive, 1977). Although there may be a preference for a positive approach, the transportation agency's campaign and other critical messages seem justifiable if they are more effective in changing behavior than positive approaches.

The attitudinal and behavioral effects of these contrasting approaches to demand modification are usually unclear because of many environmental differences, and the controversy over the choice of wording seems best settled using controlled experimentation. Unfortunately, prior research on the relative effectiveness of positive and negative appeals has not yielded conclusive findings.¹ Although most studies favor the negative appeal (Evans, Rozelle, Lasater, Dembroski and Allen, 1970; McArdle, 1972; Powell and Miller, 1967; Wheatley and Oshikawa, 1970), several have found no differences (Leventhal and Perloe, 1970; Lucas and Benson, 1929) and one has found that positively worded appeals are more effective than negatively toned ones (Yalch and MacLachlan, in review).

¹The wording or tone of a message is determined by the proportion of the message content that is positive (i.e., discusses the benefits of acquiring some good, service, or idea), compared to that which is negative (i.e., mentions the disadvantages of not acquiring the product or acquiring a substitute). Thus, a positive appeal might stress how commuting by bus is relaxing while a negative message would emphasize the tension caused by rush hour traffic jams. This definition is consistent with Lucas and Benson's (1929) view that "avoidance is the basis of negative appeal in advertising, (p. 275), but is somewhat different from that offered by Wheatley and Oshikawa (1970).

Further, a review of this literature suggests that contributions to our understanding of the message tone effect can be made in at least four areas.

One problem is that many studies used a fear appeal as the negative message (cf., Evans et al., 1970; Powell and Miller, 1967; Wheatley and Oshikawa, 1970), and it is possible that its enhancement of persuasion was caused more by the greater anxiety experienced by the subjects during their exposure to this message than to the negative tone. The finding by Sigall and Helmreich (1969) that an anxiety-arousing situation enhances persuasion supports this suggestion. Thus, it would be desirable if all the messages in an experiment on message tone produced little anxiety or, at least, that the resultant anxiety levels were equivalent for all the messages.

A second limitation of the extant studies is that most used only two levels (one positive and one negative) of message tone (an exception is Powell and Miller, 1967). Many researchers have commented on the high probability that a message characteristic will have a nonlinear relationship with the audience's attitudinal response because of enhancing and attenuating effects on the intervening stages in the communications process. For example, Ray and Wilkie (1970) postulate that there is an inverted-u curve relationship between the level of anxiety aroused by a message and the amount of attitude change caused by it. Clearly, research designs incorporating more than two levels of message tone are needed to test for the possibility that neutrally worded messages are more effective than either positive or negative appeals.

A third concern is with the measures taken to assess the effects of exposure to messages varying in tone. In line with the previously stated possibility that the effects of tone have been confounded with anxiety arousal, it would seem useful to take a measure of the acute anxiety felt during the exposure to the message. Also, the effectiveness of the tonal manipulation should be demonstrated and carryover effects to perceptions of other message and source characteristics should be checked. Further, in consideration of wide support for McGuire's (1976) multimediational model of the communication process, the audience's comprehension of the arguments used in the message as well as the attitudinal effects should be assessed. Assessment of these mediating factors is critical in formulating explanations for why one message is more persuasive than another and in developing recommendations for actual message construction.

The fourth area requiring additional research is to evaluate the moderating effects of individual personality factors on the effects of positive and negative appeals. Wheatley and Oshikawa (1970) found that a positive appeal was more effective when the audience had a high chronic anxiety level, but that a negative, anxiety-arousing appeal was more persuasive with a low chronic anxiety audience. Subsequent reanalysis of their data, however, revealed that product ownership was a more critical factor than initial anxiety in determining the effectiveness of the negative message (Wheatley, 1971). It would seem worthwhile to explore this point further using a different experimental setting as the theoretical and empirical support for the moderating effect of chronic anxiety is quite substantial (see Wheatley and Oshikawa, 1970).

These four research questions were explored in an experimental study of the effect of message tone on an audience's comprehension and acceptance of a series of arguments favoring one's living within the city limits of Seattle. The confounding effect of anxiety-arousing cues within a negative appeal was minimized by the topic selected and pilot testing of the messages. Nevertheless, a postexposure measure of felt anxiety was administered to verify that this was not an important factor. A neutral message (neither positive nor negative) was included in the design to test for curvilinear relationships, and aided recall as well as attitude questions were asked to evaluate the messages' effects. Finally, in consideration of prior theorizing about the importance of individual characteristics as moderating variables, the participants' chronic anxiety and product ownership (as measured by place of residence) were included among the personal characteristic questions.

Method

Participants

Participants in the experiment were 184 adults contacted in public places such as bowling alleys, apartment buildings, government offices, and hospitals. A non-student sample was chosen in a nonlaboratory setting in order to enhance the external validity of the results. An effort was made to recruit an equal number of persons residing in the city and its suburbs to test whether the effects of the appeals varied with the initial opinion of the audience.

Procedure

Each participant was approached and asked if he or she would cooperate in a university research project. Those agreeing were given a booklet containing a questionnaire made up of several parts and one of the four written messages. Participants were randomly assigned to treatments by sorting the booklets using a random number table before the interviewers went into the field.

The initial part of the booklet consisted of a chronic anxiety scale to assess the participants' general anxiety level. This was followed by a description of the source of the message and the message itself, all on separate pages. After reading the message, the participants answered six pages of questions. These measured their feelings while reading the message (acute anxiety), their evaluations of the message and source (manipulation checks), their understanding of what the message said (reception), their agreement with a set of statements about living in Seattle and its suburbs (post-message attitudes), and their personal characteristics. Upon completion of the booklet, the participants were thanked for their assistance and debriefed on the purpose of the study and the need to expose them to a specially prepared message.

Independent Variable

The effects of message tone were studied using three experimental (positive-neutral-negative) and one control groups, each consisting of 46 participants. Prior to the final study, a sample of city and suburban residents ranked a list of reasons for selecting where to live (e.g., quality of schools and nearness to work). The reasons ranked as most important were converted into a set of statements about the advantages of living in the city and the disadvantages of living in the suburbs. Message tone was controlled by varying the proportion of city advantages and suburban disadvantages included in the appeal.

Positive appeal. The positive appeal argued entirely for the advantages of living in Seattle. The following is an excerpt from it:

When you buy a home in Seattle you get a lot of advantages...sewers, streets, utilities, and good schools are already available. Taxes in Seattle are lower than in the suburbs, which are still in the process of establishing these vital services.

The message's other arguments included the variety of housing available in the city, better public transportation, and time saved commuting to work. It began and concluded by praising those persons still living in the city.

Negative appeal. The negative appeal argued primarily against living in the suburbs in contrast to the positive appeal's arguing for living in the city. The following excerpt was constructed to correspond to that above for the positive message:

When you buy a lot or house in the suburbs, you have to help pay for the construction of sewers, utilities, and schools...things long established in Seattle and paid for years ago. In other words, taxes are higher in the suburbs than in the city.

In addition, the message began and concluded by criticizing those individuals who had recently moved from the city to the suburbs.

Neutral appeal. A third message was developed to be neutral in tone to test for possible curvilinear effects. It was composed of statements that were less negative and less positive than those used in the other appeals. The following excerpt corresponds to those previously given:

There are both advantages and disadvantages to living in the city rather than the suburbs. Taxes are lower in the city than the suburbs because vital services like sewers, streets, utilities, and schools are already established.

In this case, the message was moderated by deleting the references to buying a home and the necessity of establishing services in the suburbs. Further, there was no mention of approval for those remaining in the city or disapproval for those moving to the suburbs.

Control appeal. Because of the after-only design of the experiment, a nonpersuasive message was developed for presentation to a no-treatment control group in order to assess the absolute persuasiveness of the three experimental messages. It contained none of the arguments used in the other messages. The following illustrates the content:

Business in Seattle has recently been clicking along at a fast pace. The products turned out in Seattle are diverse...jet aircraft, space hardware, medicine, and more recently, professional sports team. The new King County Domes Stadium is the home of the Seattle Seahawk professional football team and will soon become the home of the Seattle Mariners professional baseball team. The "Dome" has brought great excitement to the residents of Western Washington.

Although the topics discussed were related to Seattle, they were more representative of the metropolitan area than the city itself and were neutral with regard to whether it was better to live within the city limits or the suburbs.

The positive and negative messages were pilot tested on a sample of city and suburban residents to verify that they were perceived as significantly different in message tone but not on any other major message characteristic. The neutral and control messages were developed after the pilot test.

Questionnaire

Chronic anxiety. A measure of the subjects' chronic or general anxiety was taken before they were exposed to the message. Although chronic anxiety should be insensitive to situational factors, it is desirable to assess it before exposure to potentially anxiety arousing information. Initial pilot testing revealed that the 47 question, Sarason Lack of Protection scale (Adams and Sarason, 1963, p. 245) used by Wheatley and Oshikawa (1970) was overly long for a field study. Therefore, the Taylor Manifest Anxiety scale (Taylor, 1953) was used in the present study. It consists of 20 statements such as "I feel anxious about something almost all the time" and "I don't like to face a difficulty or make an important decision," each rated from definitely agree to definitely disagree using a five-point scale. The participants' responses were distributed uniformly throughout the scale with the actual range being from 24 to 99 compared to the scale's total range of 20 to 100. The reliability for the entire scale was excellent (Cronbach's alpha = .92).

Acute anxiety. After responding to the chronic anxiety test and reading one of the messages, the participants reported their acute or temporary anxiety while reading the message. This section consisted of five semantic differential items (anxious-not anxious, confident-not confident, calm-tense, threatened-unthreatened, and nervous-relaxed). The responses were summed to form the acute anxiety scale (Cronbach's alpha = .89), and used to determine if the effort to equalize the anxiety-arousing aspects of the message was successful.

Message and source evaluations. In order to verify that the message tone manipulation had been successful and to determine if tonal changes would alter the audience's perceptions of the attributed source of the message, the participants were asked to respond to two sets of semantic differential items. The first set consisted of ten bipolar adjectives describing a message. These included three on message tone (positive-negative, optimistic-pessimistic, and threatening-not threatening), two on message clarity (confusing-understandable and clear-unclear), and five on other message characteristics (strong-weak, irrelevant-relevant, exaggerated-understated, interesting-uninteresting, and constructive-destructive).

The attributed source of the message (Mr. Jim Wilson, a recognized authority in the area of urban planning and development) was evaluated in the second set of adjectives on three credibility dimensions. These were expertise (knowledgeable-unknowledgeable, competent-incompetent, experienced-inexperienced, and expert-inexpert), trustworthiness (trustworthy-untrustworthy, subjective-objective, unjust-just, and biased-unbiased), and attractiveness (not likeable-likeable, pleasant-unpleasant, similar to you-dissimilar to you, and offensive-not offensive). The three summed responses were used to determine if differences in response to variations in message tone might be attributed to changes in source credibility, and, if so, which particular attributes were affected.

Reception. It has been postulated that differences in the effectiveness of persuasive appeals might be attributed to the mediating stages in the communications process. The possibility that messages differing in tone might vary in attention maintenance or ease of comprehension was tested by having participants report the

message's position on a series of statements. For example, participants were asked whether the statement, "Taxes are lower in the suburbs than in the city," was definitely true, partly true, undecided, partly false, or definitely false according to what they thought the message said. As the correct response was always definitely true or definitely false, it was possible to score the participants by how well their recollection matched this and to sum these into an overall measure of their understanding of the message (Cronbach's alpha = .61).

Postexposure attitudes. The participants were asked to report their level of agreement (five-point scale labeled strongly agree and strongly disagree at the ends), with ten statements about the advantages and disadvantages of living in the city or the suburbs. For example, "I would recommend living in the city to a newcomer to the Northwest," was one of the statements. Several statements were not relevant to the message and served as filler items. Therefore, only six of the ten statements were summed to form the overall attitude toward living in the city (Cronbach's alpha = .82).

Personal characteristics. The final section of the questionnaire consisted of five demographic questions about the participant's age group, education level attained, place of residence, male or female, and total annual family income. These responses were used to test the success of the randomization procedure and to determine whether product ownership as determined by place of residence moderated the effects of the message tone manipulation.

Results

The results of the experimental manipulation of message tone were analyzed in terms of the effects of exposure to the message on the participants' acute anxiety, reception of the arguments, and attitudes toward living in the city.

Randomization

The randomization of the participants to treatment groups was tested using the responses to the chronic anxiety scale which was administered before any were exposed to a message and to the personal characteristics questions. An analysis of variance of the anxiety scores and chi square analysis of the personal characteristics across treatment and control groups revealed that the groups were mostly homogeneous ($p > .10$). The only variation from homogeneity was for education. As the slightly lower educational level of the neutral and control groups relative to the positive and negative message groups was not thought to be theoretically important, the assumption that the groups were randomly selected was accepted as valid.

Acute Anxiety

The possibility that the messages differed in anxiety arousal as well as tone was tested by comparing the participants' reported feelings during their exposure to it. Unexpectedly, there was a significant difference ($F = 9.4$, $df = 2,122$, $p < .001$), with the neutral message being significantly less anxiety arousing ($M = 6.4$, $n = 44$) than the positive ($M = 9.0$, $n = 40$) and negative message ($M = 9.8$, $n = 40$) according to a studentized Newman-Keuls posterior comparison of the individual cells. However, because the negative message was not perceived as more anxiety arousing than the positive message, it was felt that comparisons of the postexposure differences in these groups would be a valid test of the effect of message tone independent of anxiety arousal differences.

Message and Source Evaluation

The participants' perceptions of the message are reported in Table 1. Substantial differences were found for the tonal attributions and these were all in the predicted direction. The positive and negative message did not differ on clarity or any other characteristic. However, the neutral appeal was thought to be significantly less clear, less understandable, and less interesting than the positive and negative appeals. This may reflect the participants' difficulty in reading a message that frequently changed message tone. Because of this difference, one must be cautious in interpreting the effectiveness of the neutral message used in this study.

TABLE 1

Message Evaluations by Message Tone

Measures	Message Tone			ANOVA	
	Posi- tive	Neu- tral	Nega- tive	df	F
Tone					
Positive- Negative	4.3 ^a (44)	3.0 ^b (46)	1.7 ^c (45)	2,132	74.0***
Optimistic- Pessimistic	4.5 ^a (45)	2.7 ^b (46)	1.8 ^c (45)	2,133	91.8***
Not threatening- Threatening	4.4 ^a (44)	3.5 ^b (46)	3.3 ^b (46)	2,133	11.8***
Clarity					
Understandable- Confusing	4.3 ^a (43)	3.3 ^b (46)	4.2 ^a (46)	2,132	12.5***
Clear-Unclear	4.3 ^a (43)	3.4 ^b (46)	4.2 ^a (45)	2,131	10.5***
Other Characteristics					
Relevant- Irrelevant	3.7 ^{ab} (43)	3.3 ^a (46)	4.1 ^b (45)	2,131	6.3**
Interesting- Uninteresting	3.8 ^a (45)	3.3 ^b (46)	3.9 ^a (45)	2,133	3.8*
Strong-Weak	4.0 ^a (44)	3.3 ^b (46)	3.7 ^{ab} (45)	2,132	5.2**
Constructive- Destructive	3.5 ^a (42)	3.0 ^a (46)	3.5 ^a (45)	2,130	3.0
Understated- Exaggerated	2.4 ^a (44)	2.5 ^a (46)	2.4 ^a (44)	2,131	.1

Note: The cell means indicate how much the message was perceived as possessing the trait described by the left-most adjective. Cell sizes are in parenthesis.

abc Means in a row with different superscripts are significantly different from one another ($p < .05$).

* $p < .05$
** $p < .002$
*** $p < .001$

The source evaluations are reported in Table 2. There were no significant differences in the expertise dimension, but the source associated with the negative appeal was viewed as less trustworthy and attractive than those giving the positive and neutral messages. The neutral message source was evaluated most favorably on all three dimensions, though significantly so only in attractiveness.

TABLE 2

Source Evaluations by Message Tone

Measures	Message Tone			ANOVA	
	Posi- tive	Neu- tral	Nega- tive	df	F
Expertise	15.7 ^a (41)	16.0 ^a (45)	15.1 ^a (45)	2,128	1.1
Trustworthiness	12.5 ^a (42)	13.0 ^a (44)	11.0 ^b (44)	2,127	5.7*
Attractiveness	13.8 ^a (39)	14.9 ^b (46)	12.5 ^c (45)	2,127	10.2**

Note: High numbers indicate a more favorable evaluation. Cell sizes are in parenthesis.

abc Cell means in the same row with different superscripts are significantly different from one another ($p < .05$).

* $p < .01$
** $p < .001$

Reception and Attitude Measures

The participants' immediate recall of the message and postexposure attitudes toward living in the city are reported in the last columns of Table 3 and 4, respectively. The negative message was significantly less well recalled and also resulted in the least favorable attitude toward living in the city. In fact, in comparison with the control group, the negative appeal group gave evidence of a "boomerang" effect. That is, they shifted their attitudes in the direction opposite from the argued in the message. However, a t test of the attitudinal differences between the experimental and control groups revealed that none of the differences was significant.

Moderating Influences of Chronic Anxiety and Residence

Two audience characteristics, chronic anxiety and current place of residence, were included in the analyses of variance of the dependent measures to assess their affect on the tone-persuasion relationship. The chronic anxiety scale responses were used to divide the sample into three groups (high-moderate-low) and reported residency responses were used to determine who was or was not a resident of the city (resident-nonresident). Tables 3 and 4 report the cell means for this breakdown and Table 5 has the analyses of variance with the three messages, three levels of anxiety, and two levels of residency as the independent factors.

After considering the important influence of the participants' residency and initial anxiety, message tone had a marginally significant effect on reception and an insignificant effect on attitudes. Comparing the interaction terms showed that the moderating effect of anxiety substantially exceeded that of residency for the reception measure and slightly exceeded it for attitudes.

The prediction that a positive message would be more effective than a negative message for individuals with high anxiety was in the right direction for both reception ($M = 33.0$ vs. 27.4) and attitudes ($M = 19.7$ vs. 18.1) measures, but was statistically significant only only for the reception measure ($t = 3.53$, $df = 112$, $p < .001$).² The other prediction, that a negative

²The statistical test used is that for planned comparisons between cells (Tables 3 and 4) with the error term being estimated from the analyses of variance (Table 5). See Winer (1971, pp. 384-388).

TABLE 3

Effect of Message Tone, Chronic Anxiety,
and Residence on Message Reception

Message Tone	Nonresidents				Residents				Message Totals
	Chronic Anxiety			Total	Chronic Anxiety			Total	
	Low	Moderate	High		Low	Moderate	High		
Positive	32.8 (5)	37.0 (5)	33.2 (11)	34.0 (21)	24.7 (7)	36.4 (5)	32.7 (9)	30.9 (21)	32.4 (42)
Neutral	34.2 (6)	30.9 (9)	33.0 (5)	32.4 (20)	33.1 (10)	32.5 (10)	33.6 (5)	33.0 (25)	32.7 (45)
Negative	33.2 (6)	33.0 (7)	31.2 (5)	32.6 (18)	33.8 (6)	29.0 (7)	25.8 (12)	28.6 (25)	30.3 (43)

Note. Higher numbers indicate greater recall of the message. Cell sizes are in parenthesis.

TABLE 4

Effect of Message Tone, Chronic Anxiety,
and Residence on Postexposure Attitudes

Message Tone	Nonresidents				Residents				Message Totals
	Chronic Anxiety			Total	Chronic Anxiety			Total	
	Low	Moderate	High		Low	Moderate	High		
Positive	12.0 (5)	19.2 (5)	18.8 (11)	17.3 (21)	22.4 (7)	21.1 (8)	20.9 (9)	21.4 (24)	19.5 (45)
Neutral	16.7 (6)	16.0 (9)	17.8 (6)	16.7 (21)	24.1 (10)	22.6 (10)	20.8 (5)	22.8 (25)	20.0 (46)
Negative	15.7 (6)	15.0 (7)	12.8 (5)	14.6 (18)	22.5 (6)	20.7 (7)	20.1 (13)	20.8 (26)	18.3 (44)

Note. Higher numbers indicate more favorable attitudes toward living in city. Cell sizes are in parenthesis.

^aControl group: nonresidents, M = 14.9, n = 18; residents, M = 22.1, n = 28; total, M = 19.3, n = 46.

TABLE 5

Analyses of Variance of Reception and Attitudes
by Message Tone, Chronic Anxiety, and Residence

Source	Reception			Attitudes		
	MS	df	F	MS	df	F
Tone (A)	66.5	2	2.9*	48.1	2	1.9
Anxiety (B)	22.7	2	1.0	1.5	2	.1
Residence (C)	133.9	1	5.8**	993.8	1	40.3***
A x B	146.0	4	6.3***	22.6	4	.9
A x C	39.2	2	1.7	12.5	2	.5
B x C	8.4	2	.4	45.1	2	1.8
A x B x C	45.2	4	2.0	25.1	4	1.0
Error	23.1	112		24.7	117	

*p < .10
**p < .05
***p < .001

message would be more effective than a positive message for individuals with low anxiety, was also in the predicted direction for reception ($M = 33.5$ vs. 28.1) and attitudes ($M = 19.1$ vs. 18.1), but again the difference was only significant for the reception measure ($t = 2.75$, $df = 112$, $p < .01$).

Discussion and Conclusions

The results of this study offer further evidence that in the absence of differences in anxiety-arousing content, negatively-worded appeals are not more effective than positive messages. In fact, the data favor the use of positive appeals over negative appeals. Four issues identified in reviewing prior research (confounding of anxiety arousal with negative wording, necessity of having three levels of message tone, using multiple measures of exposure effects, and reexamining the moderator role of audience characteristics such as chronic anxiety and product ownership), were investigated by exposing participants to persuasive appeals varying in message tone but all urging individuals to live within the city limits of Seattle.

The participants' message evaluations and reported feelings during their exposure demonstrated that the positive and negative messages were successfully constructed to differ in tone without differing in anxiety arousal. However, both were significantly more anxiety arousing than the neutral message, apparently because they contained some social approval and disapproval statements. Since the neutral message was also perceived as being somewhat more confusing than the other two message, it is not clear that all postmessage differences should be attributed to the tonal differences in the messages and therefore conclusions about curvilinear effects are tenuous. The only evidence supporting the use of neutrally-worded messages was a generally more favorable evaluation of the message source compared to the positive and negative source evaluations.

One of the two major differences between the positive and negative messages was a loss of credibility for the source giving the negative appeal. The loss of trustworthiness and attractiveness suggests that an audience more easily views a negative appeal than a positive message as a persuasive attempt. This loss of credibility could be a serious concern for advertisers though there may be ways to counteract it, such as by employing highly credible sources to deliver the message. For example, the practice of having government health and safety warnings include a reference to the agency responsible would seem to be a wise strategy for enlisting public cooperation.

The other important difference was that the audience exposed to the positive message comprehended and recalled the content better than those exposed to the negative message. The fact that this led to only slight differences in postmessage attitudes probably reflects the single exposure opportunity, the message topic, and the audience. As most individuals had committed themselves to living in or out of the city, it would seem unlikely that they would easily change their opinion. With a less firmly established attitude or with repeated exposures it is possible that a significant attitude change would have been observed.

The improved reception combined with the higher source credibility evaluation offer strong support for employing only positive and neutral appeals and avoiding negative messages. Two factors argue against rigorous adoption of this recommendation, however. First, the literature review revealed that negative messages are effective when combined with other message factors, particularly anxiety-arousing information. Similarly, if very strong negative arguments are available, it is possible that

their ability to impede counter-argumentation would more than compensate for the undesirable effects of negative message tone. Under these circumstances, an advertiser would benefit by sponsoring the negative campaign.

Secondly, the effects of message tone were not uniform across audiences. The positive message was most successful with audiences initially opposed to the position being argued (nonresidents in this study), and with high anxiety individuals. Both of these groups would probably feel most threatened by the negative appeal which would result perhaps in avoidance (evidenced by the low message reception), source derogation (indicated by the unfavorable source evaluations), and counterargumentation. However, for those initially agreeing with the message's position (city residents), and low anxiety individuals, the negative appeal was as effective as the positive appeal and only slightly less effective than the neutral message. To the extent that an audience possesses any of these characteristics, an advertiser would not suffer by using a negativistic approach.

It should be noted that a reexamination of the relative importance of chronic anxiety and product ownership did not support a previous finding. Contrary to Wheatley's (1971) conclusion, a measure of chronic anxiety was a more important moderator variable than was a measure of product ownership (place of residence). There are several methodological differences between the two studies that could account for this. The present study used low anxiety-arousing appeals (all three messages yielded reported anxiety levels of less than 10 on a scale with a range from 5 to 25). The negative appeal in the original Wheatley and Oshikawa (1970) experiment, which served as the data source for Wheatley's (1971) conclusion, was designed specifically to be anxiety arousing. If all the participants are experiencing high acute anxiety because of their exposure to a threatening message than some of the communication effect attributable to individual differences in chronic anxiety could be obscured. Also, this study employed a different anxiety scale and used nonstudents in a field setting as participants which appeared to have increased the range of responses to the anxiety questions compared to the student subjects tested in a classroom in the previous study. An increase in the range of scores of an independent factor would tend to increase the likelihood that it would account for more of the variance in a dependent variable. Further, Wheatley's findings were based on attitude change scores rather than postmessage attitudes. It is possible that ownership is related to the reliability of attitudinal responses which would confound any attempt to assess the effect of the persuasive appeals. Additional research is needed to determine which of these differences account for the inconsistency in the conclusions of the two studies.

References

- Elsie B. Adams and Irwin Sarason, "Relation Between Anxiety in Children and Their Parents," Child Development, 34 (March 1963), 244-5.
- "Dummies Who Drive," Seattle Times, March 28, 1977, p. A13.
- Richard Evans, Richard Rozelle, Thomas Lasater, Theodore Dembroski and Bem Allen, "Fear Arousal, Persuasion, and Actual Versus Implied Behavioral Change," Journal of Personality and Social Psychology, 16 (October 1970), 220-7.
- Howard Leventhal and Sidney Perloe, "A Relationship Between Self-Esteem and Persuasibility," Journal of Abnormal and Social Psychology, 64 (May 1962), 385-8.

D. Lucas and C. Benson, "The Relative values of Positive and Negative Advertising Appeals as Measured by Coupons Returned," Journal of Applied Psychology, 13 (June 1929), 274-300.

J. McArdle, "Positive and Negative Communications and Subsequent Attitude and Behavior Change in Alcoholics," unpublished doctoral dissertation, University of Illinois, 1972.

William McGuire, "Some Internal Psychological Factors Influencing Consumer Choice," Journal of Consumer Research, 2 (March 1976), 302-19.

Frederic Powell and Gerald Miller, "Social Approval and Disapproval Cues in Anxiety-Arousing Communications," Speech Monographs, 34 (June 1967), 152-9.

Michael Ray and William Wilkie, "Fear: The Potential of an Appeal Neglected by Marketing," Journal of Marketing, 34 (January 1970), 54-62.

"Seattle Turns to Net TV to Promote City Living," Advertising Age, February 21, 1977, p. 20 BW.

H. Sigall and R. Helmreich, "Opinion Change as a Function of Stress and Communicator Credibility," Journal of Experimental Social Psychology, 5 (January 1969), 70-8.

Janet Taylor, "A Personality Scale of Manifest Anxiety," Journal of Abnormal and Social Psychology, 48 (April 1953), 285-90.

John Wheatley, "Marketing and the Use of Fear- or Anxiety-Arousing Appeals," Journal of Marketing, 35 (April 1971), 62-4.

_____ and Sadaomi Oshikawa, "The Relationship Between Anxiety and Positive and Negative Advertising Appeals," Journal of Marketing Research, 7 (February 1970), 85-9.

B. Winer, Statistical Principles in Experimental Design. New York: McGraw-Hill, 1971.

Richard Yalch and Douglas MacLachlan, "Persuasive Effects of Varying an Appeal's Tone, Supporting Evidence, and Source Credibility," in review.

THE RELATIVE EFFECTIVENESS OF ONE-SIDED AND TWO-SIDED COMMUNICATION
FOR MASS TRANSIT ADVERTISING

Linda L. Golden, The University of Texas at Austin
Mark I. Alpert, The University of Texas at Austin

Abstract

This research investigates the relative impact of one-sided and two-sided communication on ridership intentions, attitudes toward transit features, and advertisement specific variables. The influence of varying numbers of claims is also studied. Two-sided communication was more effective than one-sided communication in increasing copy believeability.

Introduction

For several years now the advertising prohibition against saying something negative about the advertised brand has not been strictly adhered to. Some advertisers have been "disclaiming" certain trivial characteristics of their products in conjunction with positive claims. Volkswagen, for example, used this technique to poke fun at itself while it was actually making positive claims with regard to the product.

In a disclaiming situation, an advertiser makes positive statements about characteristics that are determinants of product use, but does not claim that the product performs well on certain characteristics that are not determinants of product use. Previous research indicates that disclaiming may tend to increase the credibility of an advertisement (Settle and Golden, 1974). This may then result in a more effective advertisement.

Disclaiming in an advertisement may be viewed as providing the audience (consumers) with a two-sided argument with respect to the advertised product. For years communication research has investigated the attitudinal impact of two-sided arguments; however, the research area has not been developed in an advertising context. The focus on this research is to investigate the relative effectiveness of one and two-sided messages for mass transit advertising according to hypotheses derived from the communication research.

One-sided and Two-sided Communication Research

Communication researchers offer slightly different definitions of two-sided arguments (McGuire, 1954; Hovland, 1954; Hovland, Lumsdaine and Sheffield, 1949; Jones and Girard, 1967). Unlike other authors' definitions, Hovland's (1954) definition of two-sided arguments explicitly states that the communicator takes into account both sides of an issue, but he is himself in favor of one side. This has been described by Klapper (1949) as partial impartiality. Even though the advertiser may say something unfavorable about his brand or something favorable about the competitor, overall the advertiser will present his product as the one the consumer should buy. The advertiser is, indeed, in favor of one side. Thus, for purposes of this research Hovland's (1954) definitions are used. A one-sided argument is an argument confined to one side of an issue (Hovland, 1954).

The earliest studies in this area were designed primarily to investigate the effects of two-sided presentations. In these studies, the communicator impartially presented both sides of an argument without favoring either side. The general conclusion derived from these studies is that when one is successively

exposed to first one side and then the other of a controversial subject, the typical result is that the individual is left at approximately his/her initial position (Sims, 1938; Schanck and Goodman, 1939).

The earliest experimentation explicitly directed to the investigation of the comparative effects of one-sided and two-sided communications was conducted during World War II by Hovland, Lumsdaine and Sheffield (1949). No main effect of direct attitude change was found in this study, but there were interactions with initial favorability such that one-sided communications were more effective for those initially in favor of the conclusion and two-sided communications were more effective for those initially opposed to the conclusion. There was also a significant interaction with education such that two-sided communications were more effective for high school graduates and one-sided communications were more effective with subjects who had not graduated from high school. Later studies by Janis, Lumsdaine and Gladstone (1961), Lumsdaine and Janis (1953), and Paulson (1954) also indicated that one-sided and two-sided arguments were about equally effective over-all in producing direct attitude change.

Thistlethwaite and Kamenetsky (1955) and Thistlethwaite, Kamenetsky and Schmidt (1956) investigated the attitudinal effects of refutation of opposing arguments rather than simple mention of opposing arguments. For some groups tested, there were no significant differences between the speeches with refutation and those without. For others, the refutation speeches had more influence. The authors concluded that the speeches with mention and refutation of opposing arguments had the effect of strengthening opposing attitudes. They suggest that listeners apparently discounted the speeches with refutation as "phony" attempts to seem impartial.

All of these studies seem to suggest that mention of opposing arguments should be handled with caution. The only groups that seem more positively affected by two-sided messages were those initially opposed to the conclusion and those of higher educational levels. Even these groups did not make large changes in attitudes. Two-sided messages, however, do have a specific place in the communicator's organizational framework. They can serve to "immunize" receivers against contradictory information in later situations (Lumsdaine and Janis, 1953; McGuire, 1961; McGuire, 1962; McGuire and Papa-georgis, 1961).

Development of Hypotheses

Direct generalizations from communication research to advertising are restricted. The topics of the persuasive communications presented in the communication research were of a controversial nature. It is doubtful that the topic of many messages featured in an advertisement for consumer package goods could be considered controversial. However, advertisements for some non-traditional products or services such as birth control, welfare, and possible mass transit have topics which may be considered controversial. Further, the dependent variable in the communication literature is attitude change. The objective of advertising is to

influence, in the long-run, not only attitudes but ultimately behavior. However, given the relatively high level of educational attainment of persons likely to switch from a car to the use of mass transit (Alpert and Davies, 1975) and the relative degree of controversy surrounding mass transit compared to consumer package goods, two-sided communication is a realistic promotional tool for mass transit to explore.

The purpose of this research is to investigate the relative impact of one and two-sided communications upon ridership intentions and attitudes toward mass transportation for individuals in a sample whose demographic characteristics approximate those of "potential switchers" to mass transit identified in previous research (Alpert and Davies, 1975). Because of the relatively high level of educational attainment by persons in this sample the following hypothesis was developed:

- H₁: Two-sided messages will be more effective than one-sided messages in producing changes in attitudes toward mass transit and ridership intentions.

Since the number of positive claims included in the message may influence the effectiveness of the disclaimers, the effects of number of attributes in the copy were also tested. One hypothesis was generated to test this proposition:

- H₂: The changes produced in ridership intentions and attitudes toward mass transit by one and two-sided messages will be influenced by the number of positive attributes included in the message.

Methodology

Presentation of both one and two-sided experimental manipulations requires selection of both determinant and non-determinant attributes for mass transportation. In the two-sided manipulations, the product does not claim to possess the non-determinant attributes, but does claim to possess the determinant attributes. For transportation, determinant attributes are those attributes of a product which determine the consumer's modal choice.

Previous research in the city studied had identified determinant attributes for potential switchers. The five most determinant attributes for which the bus was rated superior to a private car were selected for use in this research. These were: economy, freedom from parking problems, freedom from repairs, low energy use per passenger, and low pollution per passenger. Given a bus's perceived superiority on these features, it is likely that advertising which asserts these as advantages might be at least believable. The selection of the non-determinant attributes required additional testing, since it was necessary that the non-determinant attributes be believable both as positive claims (one-sided) and disclaimers (two-sided). The non-determinant attributes from previous research (e.g., "ability to read," "fun to drive") could not realistically be used for both positive and negative claims, because the image of one mode was perceptually superior.

The determinancy of fifteen potentially non-determinant attributes was tested on a sample of one-hundred university students who possessed characteristics closely approximating those of potential switchers. The results indicated that the attributes colorful interior and long windows would be suitable as non-determinant attributes for both the one-sided and two-sided manipulations. These attributes were rated as relatively unimportant

transportation features for which cars and buses do not differ.

Attitudes toward both the bus and the car on these determinant and non-determinant variables (freedom from repairs, freedom from parking problems, low energy use per passenger, low pollution per passenger, economical, colorful interior, and long windows) constituted one set of dependent measures. Another set of dependent measures investigated respondents' affect toward driving a car and riding a bus. A third set of dependent measures investigated ridership intentions for specific trip purposes: to work or school and for shopping or personal business. The final set of dependent measures was designed to investigate attitudes toward the copy itself: likelihood of reading all the copy, believability, perceived quantity of information, usefulness of information, and general affect.

A pilot study was administered to a sample of 110 subjects whose characteristics closely approximated those of potential switchers, to test alternative ways of presenting the one-sided and two-sided communication formats and placement of dependent variables. A subject received one of several experimental manipulations followed by the dependent variables tentatively selected for use in the final instrument. The order of presentation was either: (1) experimental manipulation, dependent variables, media questions, or (2) experimental manipulation, media questions, dependent variables. The message formats (experimental manipulations) tested varied in their presentation of the attributes of the bus. The attributes were listed in a column below several sentences of copy, and the bus was described in one of three ways on each of the attributes. In one treatment, the bus was given a rating of either "superior" or "inferior" on the attributes. The one-sided treatment identified the bus' performance on all of the attributes as "superior." The two-sided treatment identified the bus performance on the determinant attributes as "superior" and as "inferior" on the non-determinant attributes. A second treatment followed the same general format, but replaced the adjective "superior" with "good" and "inferior" with "fair." The third treatment used check marks (✓) beside the attributes under columns labeled either "bus gives you" or "bus doesn't give you." The one-sided treatment did not contain the column "bus doesn't give you" and checked each attribute under the column labeled "bus gives you." The two-sided treatment varied in that it checked non-determinant attributes under "bus doesn't give you." The results of the pilot indicated that the use of check marks provided a slightly stronger manipulation than any of the other two treatments tested. There were no significant differences for the alternative placements of the dependent measures.

The Final Instrument

The final instrument used an after-only design with control and contained five sections. The first section presented the respondent with one of ten different experimental manipulations. The respondent could receive either a one- or a two-sided communication containing either three, four, five, six, or seven attributes. The attributes were always presented in the same order, even though the number of attributes could vary. The non-determinant attributes were always the second and third attributes presented to the respondent. The experimental manipulation was printed on heavy glossy paper on a separate page in order to simulate an advertisement as closely as possible. A cover page told the respondent that the following page contained part of an advertisement and to please read it carefully. Figure 1 exhibits a two-sided experimental manipulation with seven attributes.

FIGURE 1
Experimental Manipulation for Two-sided Message
With Seven Attributes

A lot of people are switching to the Austin city bus these days. Whether you're going to work, shopping, or just visiting, the bus will take you there. Although we're not perfect, we have many advantages. Let us tell you what the bus can and can't offer.

	Bus Gives You	Bus Doesn't Give You
Economy	✓	
Colorful interior		✓
Long windows		✓
Freedom from parking problems	✓	
Freedom from repairs	✓	
Low energy use per passenger	✓	
Low pollution per passenger	✓	

Find out for yourself what the bus can give you!

The second section of the instrument contained five questions concerning the subject's reactions to the copy. These questions were designed to ascertain the subject's likelihood of reading the copy in a magazine, the credibility of the copy, the information provided, the usefulness of the information, and the general attitude toward the copy. Responses were elicited according to a seven-point horizontal scale with one indicating the negative extreme.

The third section of the instrument obtained information regarding the subject's media habits. Information concerning the extent and nature of the subject's use of newspapers, radio, and television was elicited.

In the fourth section of the instrument, subjects were asked to indicate how likely they would be to purchase the product described in the experimental manipulation. In addition, information concerning the extent to which the subject felt the product possessed each of the seven attributes which could appear in the experimental manipulations was obtained. Subjects indicated their responses according to a seven-point horizontal scale with one representing "not at all" and seven representing "very much."

The final section of the instrument obtained demographic and personal information. Information regarding age, marital status, sex, employment status, household size, income, education, race, living situation, and number of automobiles owned was collected. On the last page of the instrument, the subject had the opportunity to request a summary of the survey results.

The final instrument was pre-tested for clarity of presentation on a sample of twenty subjects whose characteristics approximated those of potential switchers. Minor wording changes were made in the instrument as a result of the pre-test. The control group instrument was identical to that containing the experimental manipulation, except that the five questions directly regarding the advertisement (which was absent) were deleted.

Sample Selection and Administration

The criterion for the selection of subjects for the instrument was the possession of characteristics approximating those for potential switchers. In the survey area, potential switchers to mass transit had been found to be relatively younger, have smaller households, are more likely to be full-time or part-time students (although most are non-students), and they are more likely to shop and work in the downtown area than are those less likely to switch to mass transit (Alpert and Davies, 1975).

Distinct areas of a medium-sized southwestern city were identified which contained a relatively high proportion of individuals possessing the characteristics of potential switchers. An enumeration of households in these areas was obtained from Cole's Directory. In order to obtain a sample of 1,500 individuals, computer-generated random numbers were used to identify every n th person to be included in the sample frame. Only residents, not businesses, were counted when identifying potential subjects. Further, the sample was restricted to households within one-quarter mile of a current bus route, so that intention to ride the bus could be realistically measured.

Having identified the potential respondents, interviewers then began contacting by telephone. Interviewers were to ask specifically for the person whose name appeared on their calling list. Upon contact, the interviewer first gave his or her name and then requested their assistance in a consumer attitude survey being conducted by members of the Department of Marketing. When an individual agreed to participate in the study, he or she was told that they would receive the survey within a week. The respondent was instructed to please fill out the survey completely and return it at the earliest convenient time in the enclosed return envelope. Subjects were randomly assigned to treatments at the time of mailing. A letter of appreciation was included with the survey which contained the telephone number of the Department of Marketing so that the subject would have a contact point for any questions.

Statistical Analysis

The final sample consisted of 292 usable surveys returned from respondents. There were no less than 20 respondents in each cell with a maximum of 44 in the control groups.

Descriptive statistics (Veldman, 1967) confirmed that the demographic composition of the sample was highly similar to the characteristics of "potential switchers" to mass transportation derived from previous research. A second preliminary analysis performed on the data was a discriminant analysis (Veldman, 1967) to determine if respondents assigned to alternative treatments differed significantly on demographic dimensions. Three separate analyses were run: (1) comparison of respondents assigned to one-sided or two-sided treatments, (2) comparison of respondents assigned to three, four, five, six, or seven claims, and (3) comparison of respondents in each of the eleven treatment levels (including control groups). In each of these analyses, the ten demographic questions constituted the independent variables. There were no significant differences between respondents according to demographic variables for any of the above three analyses. Thus, respondents appear to have been randomly assigned to treatments on this dimension.

A final preliminary analysis was a descriptive analysis (Veldman, 1967) of the sample's ridership of the bus. Ninety-nine percent of the respondents used their car for trips to shopping or personal business. Sixty-four percent of the respondents used their car for trips to work or school; however, twenty-two percent of the respondents did not respond to this question since they did not work or go to school. Only four percent of the sample used the bus at all in the last four weeks. Thus, the sample is composed of individuals who use their car as their primary mode of transportation.

In order to compare the effectiveness of each of the experimental manipulations (advertisement treatments) against a control group, individual t-tests were performed on each of the twenty dependent variables for the respondents receiving a bus treatment. In addition, the data were submitted to two-way analysis of variance

(Veldman, 1967) for the effects of communication type (one-sided versus two-sided) and the number of claims (three, four, five, six, seven). These results are reported in Tables 1 and 2 respectively, which appear in the next section. The data from the five advertisement specific dependent variables were submitted to two-way analysis of variance (Veldman, 1967) for each dependent variable separately in order to investigate the relative effects of communication type and number of claims. These results are presented in Table 3 of the next section.

Results

Table 1 presents the results of the t-tests between the control means and the experimental manipulation means. The four most important dependent variables that can be used to evaluate the bus advertising treatments are the first four variables listed in Table 1. These measure the behavioral intentions toward the use of buses for trips to work or school (commuting) and for shopping or personal business, both over the short-run (How likely are you to ride the bus in the next month?) and "for most of your trips." As can be seen in Table 1, neither the one-sided nor the two-sided advertisement style was able to achieve any strong pattern of impact on people's behavioral intentions, which remained near the low end of the seven-point scale.

As can be seen from the cells in Table 1, there are 200 possible comparisons between the dependent measure ratings given by persons exposed to varying treatments and those in the control group. However, out of 200 possible comparisons one would expect 10 "significant" differences due to error or sampling fluctuations, using the .05 level for type-I error and two-tailed tests. There were 14 experimental rating means which were significantly different from the control mean, and of these 14, nine experimental treatment means did better than the control mean. This is hardly a strong overall pattern of change in attitudes or ridership intentions toward mass transit. Further, all the significant treatment-control comparisons for the one-sided messages were greater than the control, while only two of the

seven significant comparisons for the two-sided were greater than control. However, five of the seven significant one-sided comparisons were for non-determinant attributes (not previously salient to buses).

Table 2 indicates the significant effects obtained in the two-way analysis of variance performed on the bus variables. For four of the twenty dependent variables, communication type was shown to have a significant main effect. In all of these, one-sided communication produced a more favorable evaluation of the bus than did two-sided communication. Two of the significant dependent variables, colorful interior and long windows, were disclaimed in the two-sided treatment, so it is logical that the one-sided treatment would have more impact. The theory had hypothesized that since these features were not determinant attributes of modal choice, it would be better to "give up" some perceptions in these attributes in return for higher evaluations in terms of the determinant attributes that would be claimed. Only one attribute mentioned in the copy, low pollution per passenger, showed a significant main effect of communication type. However, the results are counter-theoretical as the one-sided was more effective than the two-sided communication. The one-sided message was also more effective for liking the bus as an alternative to the car.

The results for colorful interior and long windows probably simply indicate that respondents believed the experimental manipulation, but beyond these two variables only two dependent variables showed a significant main effect of communication type. This is one more than would be expected by chance ($\alpha = .05$), but the results indicate a stronger impact of the one-sided message. On the basis of these results, the first hypothesis which states that two-sided communication will be more effective than one-sided communication is rejected.

As further noted in Table 2, the impact of the number of claims produced significant between group variation for three of the twenty variables, which is more than would be expected by chance. A significant main effect for number of attributes was produced by: likely to ride

TABLE 1
Significant^a Mean Comparisons for Bus and Control Treatments

Dependent Variable	One-sided Number of Attributes					Control Mean	Two-sided Number of Attributes				
	3	4	5	6	7		3	4	5	6	7
Likely to ride (shopping/personal business)			1.64			1.05			1.44		
Likely to ride (work/school)						1.36					
Ride for most trips (shopping/personal business)						1.05					
Ride for most trips (work/school)						1.38					
Enjoy driving car						4.80					
Like riding bus						3.62					
Freedom from repairs (bus)						4.19					
Freedom from parking problems (bus)						4.46					
Low energy use (bus)						5.69	4.17			4.14	
Low pollution (bus)						4.80	3.58			3.52	
Economical (car)						5.07					
Colorful interior (bus)	2.95	2.80	2.81	2.68	3.00	1.81					
Long windows (bus)						3.89			2.65		
Freedom from repairs (car)						2.40					
Freedom from parking problems (car)						2.95					
Low energy use (car)						2.48			3.77		
Low pollution (car)						3.24					
Economical (car)						3.48					
Colorful interior (car)						4.27					
Long windows (car)						3.44					

^a $p < .05$

TABLE 2
Analysis of Variance for Bus Variables
(Significant^a Effects)

Dependent Variable	Treatment	MS	D.F.	F	p	Omega ²	Mean Scores
Likely to ride (shopping/ personal business)	Number of attributes	2.34	4	2.65	.03	.03	1.13 three attributes
							1.00 four attributes
							1.54 five attributes
							1.06 six attributes
							1.28 seven attributes
							5.56 three attributes
							5.49 four attributes
Enjoy driving car	Number of attributes	9.07	4	2.53	.04	.02	4.56 five attributes
							5.28 six attributes
							5.61 seven attributes
							3.17 one-sided
							2.61 two-sided
							4.62 one-sided
							3.98 two-sided
Like bus as alternate	Communication type	19.07	1	4.81	.03	.02	2.85 one-sided
							2.07 two-sided
							3.90 one-sided
Low pollution (bus)	Communication type	23.93	1	5.20	.02	.02	3.04 two-sided
							3.98 two-sided
							3.98 two-sided
Colorful interior (bus)	Communication type	32.68	1	13.43	.00	.05	2.85 one-sided
							2.07 two-sided
							3.90 one-sided
Long windows (bus)	Communication type	40.89	1	11.06	.00	.04	3.04 two-sided
							3.04 two-sided
							3.04 two-sided
Economical (car)	Number of attributes	9.78	4	2.79	.03	.03	3.75 three attributes
							4.16 four attributes
							3.31 five attributes
							4.37 six attributes
							3.45 seven attributes
							3.45 seven attributes
							3.45 seven attributes

^ap < .05

the bus for shopping or personal business within the next month, enjoy driving car, and the perceived economy of the car. No distinct pattern emerges with regard to the optimum number of attributes. It is interesting to note that the number of attributes mentioned in the advertisement about the bus influenced the perceptions of the perceived economy of the car.

There was no significant interaction between number of claims and communication type, indicating that the number of claims had no significant impact on the dependent variables -- no matter which format (one-sided vs. two-sided) was used -- given the range of claims used (three to seven). Thus, the second hypothesis, which states that the number of claims will influence the effectiveness of the communication type used, was also rejected.

In order to obtain further information on the appropriateness of particular strategies, it may be useful to consider respondents' reactions to the advertisements themselves. It may be that one reason why attitudes toward the bus were not significantly influenced by the treatments is that the advertisements themselves were at fault. Table 3 depicts the results of the two-way analysis of variance for the ratings of advertisement variables.

There were five advertisement specific dependent variables: likelihood of reading all the copy if the advertisement appeared in a magazine, perceived believability, perceived quantity of information, perceived usefulness of information, and the extent to which the respondent liked the copy. There was a significant main effect of communication type for one of the dependent variables: believability. The results tend to indicate that two-sided communication was more believable than was one-sided communication.

There was a significant main effect for the number of attributes for three of the five dependent variables: believability, quantity of information and like copy. Six attributes produced the highest mean ratings for all three dependent variables. There was also a significant interaction effect for communication type and number of attributes for the dependent variable likeability of the copy. The highest mean rating appears for the

two-sided communication with six attributes. For some numbers of claims, the one-sided communication appears to be more effective, and for other numbers of claims the two-sided communication was more effective.

Conclusions and Implications

The experimental manipulation used in this study was a piece of advertising copy administered at one point in time via a static after-only with control experimental design. This one-time exposure to a portion of an advertisement may help to account for the lack of extensive significant impacts of ridership intentions or attitudes toward the bus. However, the manipulations were strong enough to produce significantly different effects for the advertisement specific variables. In the context of the limitations of the experimental design and the product under consideration, several conclusions can be drawn from the results of this study.

Two-sided communications were no more effective in producing changes in ridership intentions or attitudes toward attributes of the bus than were one-sided communications. In fact, there was not only a lack of overall pattern of positive attitude changes, but there appears to have been a greater proportion of negative effects on specific bus features advertised for two-sided communication. It may be reasonable to speculate that the effect of advertising public transportation in a two-sided communication format to those who have the option of private transit and feel generally negative toward buses is to evoke less positive evaluations of bus features than are normally the case. It is also possible that people may be reacting against a possible attempt to influence them to use this transportation mode by rating it less positively than when they are asked (without any exposure to advertising copy) to indicate a behavioral commitment to using it. In general, attitudes toward the bus may be so negative (even for potential switchers) that even the mildest of disclaimers only serves to reinforce opposing attitudes toward the bus.

Communication research has shown that two-sided communication is likely to be more effective for higher

TABLE 3
Analysis of Variance for Ratings of Advertisement Variables
(Significant^a Effects)

Dependent Variable	Treatment	MS	D.F.	F	p	Omega ²	Mean Scores		
Believability	Communication type	23.72	1	8.18	.00	.03	4.68 one-sided		
	Number of attributes		16.75	4	5.78	.00	.07	5.31 two-sided	
								4.04 three attributes	
								4.97 four attributes	
								5.09 five attributes	
								5.64 six attributes	
								5.22 seven attributes	
Quantity of information	Number of attributes		11.57	4	4.16	.00	.05	2.64 three attributes	
								3.17 four attributes	
								3.71 five attributes	
								3.78 six attributes	
								3.69 seven attributes	
								2.30 three attributes	
								2.68 four attributes	
Like copy	Number of attributes		11.04	4	3.85	.01	.04	3.36 five attributes	
								3.38 six attributes	
								3.20 seven attributes	
								2.43 one-sided, three attributes	
								2.07 one-sided, four attributes	
								3.50 one-sided, five attributes	
	Communication type	X number of attributes		7.46	4	2.60	.04	.02	2.85 one-sided, six attributes
									3.35 one-sided, seven attributes
									2.16 two-sided, three attributes
									3.29 two-sided, four attributes
									3.22 two-sided, five attributes
									3.91 two-sided, six attributes
									3.05 two-sided, seven attributes

^ap < .05

educated persons and those initially opposed to the conclusion. Although the sample was relatively higher educated and did not tend to ride the bus, neither one- nor two-sided communication produced significant attitude or ridership intention change. It may be argued that it is unfair to expect much change in overall attitude toward the bus, given only one exposure to a partial advertisement. This is particularly a problem given the major perceived disadvantages of buses in terms of convenience, flexibility, safety from dangerous people, and other determinant attributes found in prior research in this area, and not covered in this advertisement. It may be that the key to changes in ridership intentions and attitudes toward the bus lies in changes in the system and not in the nature of the promotional message used for this product. Attitudes may be so negative that a product change may be required to shift them, because the product may not offer significant advantages (possibly it offers significant disadvantages) over the most popular mode of transportation: the car.

Two-sided communication does appear to produce more believable copy than does one-sided communication. However, the believability is such that it does not make attitudes more positive and may in fact cause attitudes to move in a negative direction toward the attributes for which positive claims are made.

The number of attributes described in the copy tended to be more critical to the effectiveness of the advertisement for this research than did the communication format employed. For attitudinal variables toward the bus, there was no clear pattern of the optimum number of attributes. However, this research indicated that incorporating six attributes in the copy produced the highest ratings on advertisement related variables. Consumers tended to believe the advertisement more, perceived more information and liked the copy better when it contained six attributes. In addition, liking the copy may be a function of both the number of attributes and the communication type. Clearly, the operational

relevance of this information depends upon the objectives to be achieved by the advertisement.

In order to more completely investigate the relative impact of one- and two-sided advertising, this research needs to be extended to more products, more independent and dependent variables, and to potentially stronger but realistic manipulations, over time. From the results of this research it appears that two-sided communication may be less effective than one-sided communication in producing positive attitude change toward product features. However, two-sided messages appear no more or less effective in influencing ridership intentions. Further testing is required with a myriad of situations and variables before any definitive relative effectiveness information can be discerned. What is likely to be most effective for mass transit is changes in the system along critical determinant attributes.

References

- M. Alpert and S. Davies, The Marketing of Public Transportation: Method and Application, Research Report 19, Council for Advanced Transportation Studies (The University of Texas at Austin, 1975).
- L. Golden, "Attribution Theory Implications for Advertisement Claim Credibility," Journal of Marketing Research, 1(1977), 115-117.
- C. Hovland, "Effects of the Mass Media of Communication," in G. Lindzey, ed. Handbook of Social Psychology Vol. 2 (Cambridge: Addison-Wesley, 1954).
- C. Hovland, A. Lumsdaine, and F. Sheffield, Experiments in Mass Communication: Studies in Social Psychology in World War II Vol. 3 (Princeton: Princeton University Press, 1949).

- I. Janis, A. Lumsdaine and A. Gladstone, "Effects of Preparatory Communications on Reactions to Subsequent News Events," Public Opinion Quarterly, 2(1951), 487-518.
- E. Jones and H. Gerard, Foundations of Social Psychology (New York: John Wiley and Sons, 1967).
- J. Klapper, The Effects of Mass Media (New York: Columbia University Bureau of Applied Social Research, 1949).
- A Lumsdaine and I. Janis, "Resistance to 'Counterpropaganda' Produced by One-Sided and Two-Sided 'Propaganda' Presentations," Public Opinion Quarterly, 17(1953), 311-318.
- W. McGuire, "The Nature of Attitudes and Attitude Change," in G. Lindzey, ed. Handbook of Social Psychology Vol. 3 (Cambridge: Addison-Wesley, 1954), 136-314.
- W. McGuire, "Persistence of the Resistance to Persuasion Induced by Various Types of Prior Belief Defenses," Journal of Abnormal Social Psychology, 64(1962), 241-248.
- W. McGuire and D. Papageorgis, "The Relative Efficacy of Various Types of Prior Belief-defense in Producing Immunity Against Persuasion," Journal of Abnormal Social Psychology, 62(1961), 327-337.
- S. Paulson, "The Effects of Prestige of Speaker and Acknowledgement of Opposing Arguments on Audience Retention and Shift of Opinion," Speech Monographs, 21(1954), 267-271.
- R. Schanck and C. Goodman, "Reactions to Propaganda on Both Sides of a Controversial Issue," Public Opinion Quarterly, 3(1939), 107-112.
- R. Settle and L. Golden, "Attribution Theory and Advertiser Credibility," Journal of Marketing Research, 1(1974), 181-185.
- V. Sims, "Factors Influencing Attitude Toward the TVA," Journal of Abnormal Social Psychology, 33(1938), 34-36.
- D. Thistlethwaite and J. Kamenetsky, "Attitude Change through Refutation and Elaboration of Audience Counterarguments," Journal of Abnormal and Social Psychology, 51(1955), 3-12.
- D. Thistlethwaite, J. Kamenetsky and H. Schmidt, "Factors Influencing Attitude Change through Refutative Communication," Speech Monographs, 23(1956), 14-25.
- D. Veldman, Fortran Programming for the Behavioral Sciences (New York: Holt, Rinehart and Winston, 1967).

COLA PREFERENCES: DISGUISED TASTE VS. BRAND EVALUATIONS

Thomas J. Stanley,¹ Georgia State University

Abstract

This study was designed to determine if preferences for colas based on taste evaluations could explain overall cola brand preferences. In 25 of 60 cases studied, subjects' taste preferences were significantly related to their brand preferences. Thus, evidence supports the contention that taste is an important evaluative criterion that underlies consumer cola brand preferences.

Introduction

Recent studies (Lane, Zychowski, and Lelli, 1975; Thumin, 1962) have indicated that subjects can distinguish between popular brands of colas; however, these studies have not determined if the taste of colas is an important factor in explaining overall cola preferences. Several scholars (Posner, 1974; and Warne, 1962) suggest that the physical differences of products such as colas are usually unimportant factors in explaining consumer preferences. Social critics conclude that differentiation of cola brands via promotion of taste are nothing more than "foolish puffing." Governmental agencies also frown on what is called puffing ads, which promote in their words "pseudo-differences" (Backman, 1968; Posner, 1974; and Warne, 1962). In actuality, the cola industry has steered away from promoting direct taste comparisons, even though consumers disagree with the idea that product differences between competing brands of most products are insignificant and unimportant to them (Barksdale and Darden, 1972). Historically, the majority of cola advertisements have not emphasized taste (Block, 1975), but have attempted to develop strong brand images for their products by associating them with emotionally loaded themes such as brotherhood, nostalgia, and reference groups. Critics of advertising argue against the use of these conditioning methods.

Cola producers would have a strong argument as well as motivation for the use of advertising which emphasizes taste differences if they could demonstrate that taste is an important factor in explaining cola preferences. The success of future product as well as promotional developments in this context may hinge upon whether or not taste is an important factor in explaining overall brand preferences for colas (Block, 1975). This paper attempts to determine if taste preference evaluations can explain overall brand preferences for colas.

Method

Subjects

Sixty students from a state university participated in this study. There were 41 male and 19 female subjects who ranged in age from 19 to 27 years. Eleven of the original 71 volunteers did not participate in the study; five subjects indicated that they were not familiar enough with colas to evaluate them; and the other six

withdrew before completing their evaluations.

Procedure

The stimuli used in this study were seven cola drinks including: (1) Adirondack Cola; (2) Canada Dry Cola; (3) Coca-Cola; (4) Cott Cola; (5) Pepsi Cola; (6) Shasta Cola; and (7) Tab. Adirondack and Cott are regional soft drinks while the others are well known nationally. Shasta and Tab were the only diet colas used in the experiment.

Each subject evaluated two separate sets of stimuli. Subjects were first asked to indicate their preferences for the colas, in a blind, paired comparison taste test. Each subject rated all 21 pairs and indicated a preferred cola in each pair. No information about the identity of the cola was given the subjects prior to or during this part of the experiment. Approximately four ounces of each cola was given to each subject in opaque plastic cups. The colas were stored at approximately 5°C prior to their distribution. Salt free flour crackers were also distributed and subjects were instructed to consume a bit of the crackers after tasting each cola pair. The position of the colas in every pair as well as the order that pairs were presented in was randomized.

Following the completion of the taste evaluations each subject made preference evaluations on a set of 21 randomly ordered pairs of the names of seven cola brands. The brand names were those of colas tasted in the first step of the experiment. The responses to cola names (tastes) were used to construct a brand (taste) preference scale.

The Model

Both the taste and brand preference data were separately analyzed by the MD-PREF scaling algorithm. This program transforms subjects' paired comparison preference data matrices into rank orders or first score matrices. The Eckart-Young factor analytic procedure then decomposes the ranked data and generates a group stimulus space or matrix and termini of subject vectors (ideal points) that best agree in a least squares sense with the matrices of subjects' scaled preference values. In other words, MD-PREF determines the stimulus coordinate values and termini of subject preference vectors from the set of preference judgments that will most closely agree with the preference scale values for each subject (Green and Rao, 1972; and Green and Wind, 1973).

The output from MD-PREF includes the stimuli which are represented as points and subjects as vectors in the same space. A subject's rank preference ordering is depicted as a vector drawn through the origin of the stimulus space to the subject's ideal point. The correlation between the vector and a particular coordinate axis indicates how well a subject's preference ordering can be explained by the positions of the stimuli along that axis. The farther the subject's ideal point is from the origin the better the stimulus configuration can explain a subject's preference ordering (Carroll, 1972).

Rationale

Today, the average cola consumer must decide what cola

¹The author is indebted to Dr. Ralph L. Day, Professor of Business, Indiana University, for his helpful suggestions.

to purchase from a large set of offerings. From the results of previous cola taste experiments (Thumin, 1962; Lane, Zychowski, and Lelli, 1975) it is unclear if taste plays a role in explaining cola preferences. In this study subjects were asked to evaluate seven colas given in 21 pairs. This number (seven) may approximate the typical number of cola brands generally offered in the market area where the study was conducted. However, the possibility exists that subjects who tasted 21 successive pairs of colas encountered taste adaptation. Some amount of perceived adaptation may typify the actual market conditions under which consumers must make such decisions. Also adaptation would make it more difficult to reject the null hypothesis that no relation exists between cola preferences based on taste evaluations and those based on brand preference evaluations. Rejection of the null hypothesis under such conditions would provide a stronger argument that taste preferences are generally important in explaining brand preferences.

Both the taste and brand preference data were separately analyzed by the MD-PREF multidimensional scaling algorithm. The resulting preference orderings (taste preference and brand preference scales) for each individual were correlated to determine the extent to which taste evaluations could explain evaluations based on brand names. The multidimensional scaling method was used instead of more traditional scaling methods for several reasons. An important aspect of multidimensional scaling methods lies in their ability to accurately reproduce, in the form of a multiple dimension geometric space, preference as well as perceptual configurations (Dubois, 1973; Green and Carmone, 1970; and Stanley, 1976, a, b).

Results

Tables 1 and 2 illustrate the stimulus spaces from the MD-PREF analyses of taste and brand preference evaluations. Two factors explained 70 percent of the subjects' taste judgments and 78 percent of the preference evaluations based on brands. Factor one explained slightly more than 50 percent in both cases. In both cases the third factor added little to the explanation of preference judgments and was therefore excluded from further analysis.

Tables 1 and 2 demonstrate that in regard to factor one high negative scores are associated with the diet colas, namely Shasta and Tab, while high positive scores are noted for Pepsi and Coca-Cola. The direction of subjects' ideal points suggest that positive values are indicative of more preferred colas than those with negative values.

TABLE 1
Stimulus Matrix Generated
From Cola Taste Evaluations

Stimulus	Factor	
	I	II
Adirondack Cola	.0600	.2623
Canada Dry Cola	.1722	-.3947
Coca Cola	.2998	.0552
Cott Cola	.2700	.5126
Pepsi Cola	.3595	-.3706
Shasta Diet Cola	-.5890	.3979
Tab	-.5725	-.4626

TABLE 2
Stimulus Matrix Generated
From Cola Brand Evaluations

Stimulus	Factor	
	I	II
Adirondack Cola	-.3875	-.4339
Canada Dry Cola	.1221	.0400
Coca Cola	.5191	.1699
Cott Cola	-.0203	-.2856
Pepsi Cola	.5140	.0408
Shasta Diet Cola	-.2697	-.3078
Tab	-.4777	.7767

Factor one in both preference spaces indicates that subjects consistently evaluated Coke and Pepsi as distinctly more preferred over the diet colas. Therefore, the evaluations of Coke and Pepsi and the diet colas would make a substantial contribution to high positive correlations that may be found between taste and brand preferences. Without the assistance of other measures such as perceptual and corresponding property scales, no attempt to label the factors generated from an internal preference analysis should be made.

Subject Correlations

Two rank orders (individual taste and brand preferences) of the stimuli were computed from the two factor MD-PREF outputs for each of the 60 subjects along their preference vectors. In turn, 60 Spearman rank correlations were computed between the ranks assigned by each subject to colas on the basis of taste and brand evaluations (see Table 3).

Twenty-five of the 60 correlations were positive and significantly greater than zero, at the $p < .05$ level. The chance expectation at the $p < .05$ level is that only three of the 60 subjects will have a significant rank order correlation. As determined by the Poisson approximation to the binomial distribution, the probability of finding 25 or more of the 60 significant correlations where no actual relationship exists between taste and brand preference for any of the subjects is less than .001, where $n = 60$, $p = .05$, and $M = np = 3.0$.

Discussion

The results of this study indicate that a significant relationship exists between cola preferences based on taste evaluations and brand preferences. Therefore, a possibility exists that cola brand preferences could be predicted from taste evaluations. Thus, cola producers may wish to consider placing more emphasis on promoting and developing unique flavor characteristics for their products. By demonstrating that variations in brand preference can be explained by taste evaluations cola producers can counter social critics and government agencies who condemn the use of product differentiation themes in cola promotions.

Multidimensional scaling techniques appear to be a useful tool in measuring consumer taste as well as brand preferences. Future studies should investigate the perceptual properties that explain taste and corresponding brand preferences. By being able to understand the variations in the saliences consumers attached to the dimension of taste, cola producers will be in an excellent

position to develop products which are more closely tailored to the needs of various market segments.

TABLE 3
Spearman Rank Correlations Between Subject Taste and Brand Preference Orderings¹

Subject	Correlations ²	Subject	Correlations	Subject	Correlations
1	.036	21	.946*	41	.429
2	.821*	22	.589	42	.929*
3	.857*	23	.839*	43	.607
4	.411	24	.571	44	.536
5	.643	25	.536	45	-.286
6	-.357	26	-.036	46	.536
7	.893*	27	.893*	47	.857*
8	.429	28	.893*	48	.857*
9	.929*	29	.750*	49	.607
10	.857*	30	.286	50	.143
11	.786*	31	-.250	51	.750*
12	.607	32	.786*	52	.607
13	.250	33	.357	53	.679
14	-.286	34	.714*	54	.750*
15	.607	35	.857*	55	.357
16	.429	36	.571	56	.714*
17	.607	37	.857*	57	.857*
18	-.357	38	.607	58	.804*
19	.750*	39	.321	59	.607
20	.857*	40	.607	60	.643

¹Rank preference orderings were computed from each subject's preference vectors.

²Indicates (*) significant correlation, $p < .05 = .714$

References

- Jules Backman, "Is Advertising Wasteful?" Journal of Marketing, 32 (January, 1968), 2-8.
- Hiram C. Barksdale and William R. Darden, "Consumer Attitudes Toward Marketing and Consumerism," Journal of Marketing, 36 (October, 1972), 28-35.
- Maurine Block, "Pepsi Dares Coke in Texas Taste Text," Advertising Age, (April 14, 1975)
- J. D. Carroll, "Individual Differences and Multidimensional Scaling," in R. N. Shepard, A. K. Romney, and S. B. Nerlove, (eds.), Multidimensional Scaling Volume I (New York: Seminar Press, 1972), 105-155.
- B. Dubois, "Ideal Point Versus Attribute Models of Brand Preference: A Comparison of Predictive Validity," in Mary Jan Schlinger (ed.), Advances in Consumer Research, Vol. 2 (Chicago: Association for Consumer Research, 1975), 321-333.
- Paul E. Green and Vithala Rao, Applied Multidimensional Scaling (New York: Holt, Rinehart and Winston, Inc., 1972), 7-8, 212-214.

Paul E. Green and Yoram Wind, Multiattribute Decisions in Marketing (Hinsdale, Illinois: The Dryden Press, 1973), 108-109.

Paul E. Green and Frank J. Carmone, Multidimensional Scaling (Boston: Allyn and Bacon, 1970), 82-84.

Sam H. Lane, James Zychowski, and Kenneth Lelli, "Cola and Diet Cola Identification and Level of Cola Consumption," Journal of Applied Psychology, 60 (1975), 278-279.

Richard Posner, "Advertising and Product Differentiation," in J. S. Wright and J. E. Mertes (eds.), Advertising's Role in Society (St. Paul, Minnesota: West Publishing Co., 1974), 44-46.

Thomas J. Stanley and Murphy A. Sewall, "Image Inputs to a Probabilistic Model: Predicting Retail Potential," Journal of Marketing, 40 (July, 1976), 48-53. (a)

Thomas J. Stanley, "Predicting Retail Trade via Chain Preference Measures," a paper presented at the 1976 Conference of the American Psychological Association, September 6, 1976, Washington, D. C. (b)

F. J. Thumin, "Identification of Cola Beverages," Journal of Applied Psychology, 46 (1962), 358-360.

Colston E. Warne, "Advertising--A Critic's View," Journal of Marketing, 26 (October, 1962), 10-14.

NONMETRIC UNIDIMENSIONAL SCALING OF CONSUMER
PREFERENCES FOR PROPOSED PRODUCT DESIGNS¹

Murphy A. Sewall, State University of New York at Albany

Abstract

Multidimensional scaling is an increasingly popular device for analyzing consumer preferences. However, in some instances a one dimensional scale may provide an adequate basis for decision-making. This paper describes a procedure for unidimensional scaling of rating data and presents empirical confirmation of validity based upon an independent data set.

Introduction

Recent years have seen the development and application of increasingly sophisticated multidimensional scaling techniques for examining consumer perceptions and preferences (Green and Rao, 1972). The need to examine two or more components, or dimensions, may be necessary to understanding the mental process underlying relative perceptions and preferences in some cases. However, two arguments can be made favoring consideration of the alternative of unidimensional scaling.

From the point of view of the decision-maker, the relevant question is usually: "Which alternative(s) among a set of proposals is most preferred by potential customers?" Such a question is clearly a request for a ranking at minimum, and a ranking that indicates the relative differences between closely ranked proposals if possible. If a scale of only one dimension can fairly represent consumer preferences, such a scale is clearly consistent with the needs of the user (and sponsor) of the research.

Also, a simple explanation is preferable to a complicated one that offers little additional information. Multidimensional solutions are often difficult to interpret, difficult to explain, and hard to understand. Their implications for decision-making may be so difficult to determine that their practical value may be much less than their elegance. While much of the "real world" may be inherently too complex for anything less than a multidimensional solution, it could be a serious mistake to assume, without examination, that more than one dimension is needed to describe or explain a phenomenon.

This paper represents results of a survey of consumer reactions to 36 proposed bed linen designs. The analysis indicates that a nonmetric unidimensional scaling procedure does provide an adequate representation of relative group preference for this particular problem. The scale predicts the relative preferences of an independent ("holdout") sample of consumers with a high degree of accuracy.

The Problem

The 36 proposed designs evaluated in this study were supplied by a major seller of bed coverings (and a variety of other consumer products). The competition for the 650-700 million dollar market for bed linens has

grown increasingly intense. Over thirty name fashion designers are currently producing patterns for competing marketers of these products. In spite of the proliferation of designs, generic demand for some items in the line is actually declining, for example, approximately three quarters of sheets are sold at sale prices (Reif, 1977).

Selecting which patterns to introduce from among those proposed by designers has been an important problem for product managers ever since the advent of "designer" bed linen a decade ago. The increasing intensity in today's market places an even greater premium on preference information obtained from pre-tests of proposed designs on consumers.

The number of new patterns proposed for each selling season is typically twenty or more. Management would like to evaluate as many choices as possible since the number of potentially strong entries identified is believed to increase with the number of ideas examined. A researcher would prefer to keep the number of alternatives small in order to keep the survey task manageable. Requesting too many responses from individual consumers is likely to increase the frequencies of refusals to participate, refusals to provide complete responses, and rote responses which do not reflect the true preferences of the subjects.

Survey Design

Most preference scaling techniques are based on expressions of subjects' preferences for one of each possible pairing of all alternatives. Unfortunately, the number of pairs to be evaluated expands geometrically as the number of alternatives increase.

It is unreasonable to expect consumers to respond to the large number of paired comparisons that would be required to evaluate the number of bed linen designs that management wishes to consider. Hence, it is necessary to utilize a data collection technique which requires far fewer responses from each subject. A rating scale approach was selected because only one response per pattern is required, and rating also presents respondents with a readily comprehensible evaluation task.

The thirty-six patterns were photographed on a bed, without head board or foot board, against a plain white background and arranged on five display boards, each of which represented a designer "collection." Subjects were first asked to examine each collection as a whole. After they studied all five collections and were given an opportunity to consider their preferences over the entire set of patterns, respondents were asked to rate each pattern on a five point scale:

If this pattern were offered for sale -

1. I would definitely buy.
2. I probably would buy.
3. I might or might not buy.
4. I probably would not buy.

¹This project was supported by a Fortune 500 firm as part of a SUNY Albany MBA field project. The sponsor wishes to remain anonymous for proprietary reasons.

5. I definitely would not buy.

The survey also included a list of demographic questions (age, income of household, and so forth), questions about the number of bedrooms in the subject's home and the style(s) of furniture in each, and other questions about whether various attributes of the linen or the marketing firm were important to their purchase decisions.

Data were obtained by shopper intercept interviews at shopping centers in eight metropolitan areas (two in the Northeast, one in the Southeast, three in the Midwest, and two on the West Coast). A total of 399 interviews were completed (approximately 50 in each city).

Data Analysis Procedure

The use of rating scales in marketing research is quite common, and has been for some time. However, quantitative analysis of this type of data often involves statistical calculations which are used to draw inferences which are conceptually inappropriate (Adams, Fargot, and Robinson, 1965). The usual errors are the implicit assumptions that:

1. Adjacent response categories are equidistant (that is the data are interval scaled).
2. These intervals are constant across subjects.

Rating Scale Analysis

The most common approach to analyzing five-point intention to purchase ratings consists of ranking the stimuli in order of the fraction assigning one of the first two (would buy) ratings (Taylor, Houlahan and Gabriel, 1975). This procedure is subject to several criticisms. The calculations discard information by reducing the five-point ratings to a two-point scale. There is an implicit assumption that all subjects mean the same thing by "probably would buy." And, the analysis does not provide a statistical basis for determining how much more desirable a highly rated pattern is over one with a lower rank.

One paper that explicitly recognized the problem of assuming too much about rating data (Harris, 1964) is subject to similar criticism for using an arbitrary rule for determining which values should be counted (the top three categories in a 21 point scale which actually produced a forced ranking of the alternatives). The analysis in the study also assumes homogeneity across subjects in their interpretation of the range covered by the scale. Attempts to validate the scaling against a "sales index" for two types of patterns produced widely divergent results. While the explanation that is given for the disparity appears plausible, it is qualitative. The possibility that the one high correlation between the scale and the sales index was due to chance cannot be eliminated.

A procedure has existed for over 25 years for scaling rating data without assuming equal distances between categories (Edwards, 1952). The method of successive intervals does assume that subjects make the same evaluation of each response category (subject to some error in making individual judgments about the assignments of stimuli to categories). That is, scaling by this method assumes, for example, that all responses of "I probably would buy" represent the same probability of purchase for all subjects.

The method proposed in this paper is less restrictive in its assumptions about the original data than the method of successive intervals. It is also easy to use the scale reported below to verify goodness-of-fit to the original data or to predict responses of subjects from a

different sample, although it is possible to perform these calculations with the method of successive intervals.

Plans were made to scale the data for this study by the method of successive intervals for comparative purposes. Unfortunately, a limitation of the successive intervals approach aborted these plans. The frequency distributions for four of the thirty-six designs were found to be badly skewed. More than half of the subjects stated that they definitely would not buy these four designs. The method of successive intervals fails if the median occurs in an end point of the rating scale for any stimulus.

The skew in some of the frequency distributions for the pattern ratings is hardly surprising. The patterns are, after all, new product ideas, and the high consumer rejections rate for new product ideas is well known. The fact that skewed rating data is to be expected in this type of consumer research is an obvious indication that metric data analyses techniques are inappropriate.

Derived Paired Comparisons

The approach taken in this study is, in concept, one of restructuring the data into a form that has previously been solved. Specifically, each subject's ratings were converted into a set of pairwise preferences. The notion of a "derived paired comparison matrix" has been considered before. Neidell (1972) recommends derived measures for large stimulus sets. Green and Rao (1972, p 25) propose derived paired comparisons from multi-attribute rating profiles. Five-point intention to purchase ratings may be viewed as sorting n stimuli into k classes (Rao, 1971).

As proposed here, derived paired comparisons for each subject are obtained by comparing the two ratings for each possible pair and assuming the subject prefers the design toward which he has expressed the more favorable purchase intention (Reynolds, 1966). Tied ratings are treated as missing data. That is, subsequent calculations of the fraction of respondents expressing a preference for one pattern over another are based only on that portion of the sample that gave a clearly more favorable rating to one pattern. The alternative approach would be to treat ties as a value of one-half for each pattern in a pair (Benson, 1962). Computations were made for both assumptions about ties. Results were equivalent to two decimal places. The missing data approach is favored by the sponsoring management since subjects who expressed a greater degree of distinction between the patterns produce fewer tied comparisons and are, therefore, given greater weight in the resulting group scale. It is assumed that subjects who gave many patterns the same rating are relatively indifferent to which designs are marketed. There is a more sophisticated, but computationally complex, method for treating tied (no preference) data (Draper, Hunter and Tierney, 1969). Application of this procedure was not deemed appropriate, from a cost-benefit point of view, for this stage of the analysis.

Only two assumptions need to be made about the data in order to justify this formulation of derived paired comparison matrices theoretically:

1. Each subject applied a consistent underlying scale to his evaluations of all patterns presented.
2. This underlying scale is at least ordinal.

The first assumption implies that the subject is consistent with himself (but not necessarily with other subjects). The second assumes that his ratings are consistent with the stated rating categories. Since

the subjects studied all the patterns and responded to opinion questions about groupings of them (the five "collections") before rating the patterns individually, the likelihood that they formulated personal scales which remained nearly fixed over the 36 ratings requested is plausible.

Unidimensional Scaling

The derived paired comparison preference matrices are summarized (fraction of subjects preferring the column stimulus to the row stimulus for all columns and rows) and subjected to Thurstone scaling (Case V). The theory and method of this procedure was originally published fifty years ago (Thurstone, 1927) and has subsequently been treated in a wide variety of papers and texts (Mosteller, 1951a and 1951b, among many others). Thus, Thurstone V scaling is taken as "previously solved" insofar as mechanics and conceptual justification are concerned.

The scaling of stated intention to purchase data by this procedure may be viewed as an unusual application by some. A consumer's evaluation of his purchase intention is likely to involve a considerable number of considerations. Hence, one might think that purchase intention is far too complex a concept for Thurstone's procedure. However, at the outset, Thurstone noted that:

"It is not necessary to limit psychological analysis to stimuli which have intensity or magnitude as their principal attributes. For example, a series of handwriting specimens may be arranged in a continuum on the basis of general excellence (Thurstone, 1927, p. 369)."

A judgment about which of two handwriting specimens is "more excellent" does not seem to be altogether different from an evaluation of which of two bed linen designs is more desirable.

Fortunately, a statistical text exists (Mosteller, 1951b) for the adequacy of applying Thurstone V assumptions to a particular set of data. Numerically, the statistic involves a chi-square calculation based on arc sine transforms of the square roots of the predicted and observed percentages of subjects preferring one of each possible pair of stimuli. Thus, an empirical basis for arguing that an application of Thurstone V scaling does not violate the procedure's assumptions is available. This statistical test, and some others, may also be applied to predictions of the fraction of respondents preferring one pattern over another from a sample of subjects other than those used to develop the scale.

Results

Ratings for the 399 respondents to the survey were randomly sorted into two groups. These groups are labeled 'A' (195 subjects) and 'B' (204 subjects) in the discussion and tables below.

Goodness-of-Fit

Derived paired comparison matrices, and subsequently Thurstone scales, were formulated for each group. The scales were then used to predict the fraction of respondents preferring one of each of the 630 possible pairings of patterns (in matrix form, the fraction preferring the column designs over the row designs for the upper half matrix with missing diagonal).

Tables 1 to 3 each contain four statistical results. Since the data are randomly divided into two subsamples, it is possible to use each as a "holdout sample" to test the predictive power of a scale derived from the other.

The same statistics also may be calculated in the usual manner - using the subsample which produced a scale as the observed data (a less rigorous goodness-of-fit test than one based on an independent sample).

Table 1 contains statistics for Mosteller's (1951b) goodness-of-fit. The very low chi-square values indicate a close correspondence between scale predictions and observed proportions.

TABLE 1
Mosteller's Goodness-of-Fit Test

Scale From Sample	Observations From Sample	χ^2 d.f. 595	Tail Probability
A (n = 195)	A	.204	.983
A	B	.664	.972
B (n = 204)	B	.202	.983
B	A	.671	.972

Table 2 represents an alternative (and perhaps more familiar) perspective of the goodness-of-fit. This table contains correlation statistics between predicted and observed fractions of subjects preferring one of each pair of designs (630 predictions). The values are surprisingly high. The null hypothesis that the true correlations are zero is rejected at any reasonable significance level. It does not appear that multidimensional scaling would be likely to provide additional insight into the samples' aggregate preferences for bed linen designs.

TABLE 2
Product Moment Correlations

Scale From Sample	Observations From Sample	Correlation
A (n = 195)	A	.978
A	B	.920
B (n = 204)	B	.975
B	A	.921

Desirable Patterns

Table 3 contains statistics for means, standard deviations, and mean absolute deviations of the 630 prediction errors. The ninety-five percent confidence interval (based upon across sample error deviations) is approximately twelve percent. That is, if pattern 'A' is predicted to be preferred over 'B' by 62 percent or greater, then the likelihood that 'B' is actually preferred to 'A' is five percent or less.

TABLE 3
Prediction Errors

Scale From Sample	Observations From Sample	Mean Error	Mean Absolute Deviation	Standard Deviation
A (n = 195)	A	-.002	.025	.034
A	B	.006	.049	.061
B (n = 204)	B	-.003	.024	.033
B	A	-.010	.050	.063

Since management seeks above average designs, a decision-rule to classify a pattern as desirable if it has a predicted preference of 62 percent or greater compared to an "average pattern" (the mean preference scale value for all 36 patterns) was recommended. Seven of the 36 patterns in this study were identified as satisfying this decision-rule.

Conclusion

After the fact, it does not seem surprising that a scale of one dimension provides an excellent fit to the preferences derived from consumer intention to purchase ratings. A recent article discussing applications of conjoint analysis (Hupfer, 1977) points out that attempts to break down preferences for product choices into component attributes and benefits are probably inappropriate when relatively little conscious thought is given to the purchase. It seems plausible that consumers' analyses of cost-benefit trade-offs occur in choosing whether or not to buy new bed linen, and if so, in what fabric (muslin or percale) and style (flat or fitted) and for what purpose (beds, walls, windows, or furniture) (Reif, 1977). The choice of pattern from among the many competing within chosen fabric and price groups may indeed be made with little conscious analysis.

It is likely that nonmetric unidimensional scaling will prove appropriate for similar applications in other product areas. Certainly, in instances where brand competition is characterized by proliferation of "fashion designs," results similar to those obtained in this study may be expected.

Unidimensional scaling also appears to have merit as a preliminary to multidimensional scaling. When direct paired comparison data is obtained, Thurstone V scaling is straight-forward. It seems surprising that applications of Thurstone scaling rarely appear in recent marketing literature.

It also seems puzzling that scaling by the method of successive intervals (Edwards, 1952) has received virtually no attention from marketers. The method is computationally simple (particularly when programmed for a computer). One would expect that this method would be appropriate for the data collected in many studies (Reynolds, 19-6, for example). The nonmetric unidimensional scaling procedure described here may prove to be more general, and equivalent to successive intervals when both methods are appropriate, but this assertion remains to be verified.

References

- Adams, E. W., R. F. Fargot and R. E. Robinson. "A Theory of Appropriate Statistics," Psychometrika, 30 (June 1965), 99-127.
- Benson, Purnell H. "A Short Method for a Distribution of Consumer Preferences," Journal of Applied Psychology, 46 (October 1962), 307-13.
- Draper, N. R., W. G. Hunter and D. E. Tierney. "Analyzing Paired Comparison Tests," Journal of Marketing Research, 6 (November 1969), 477-80.
- Edwards, A. L. "The Scaling of Stimuli by the Method of Successive Intervals," Journal of Applied Psychology, 36 (April 1952), 118-22.
- Green, Paul E. and Vithala R. Rao. Applied Multidimensional Scaling. New York: Holt, Rinehart and Winston, Inc., 1972.
- Harris, Douglas, "Predicting Consumer Reactions to Product Designs," Journal of Advertising Research, 4 (June 1964), 34-37.
- Hupfer, Herbert. "Techniques Useful in Planning New Products, Especially Costly Ones," Marketing News, 10 January 28, 1977) 10 ff.
- Mosteller, Frederick. "Remarks on the Method of Paired

Comparisons: I. Least Squares Solution Assuming Equal Standard Deviations and Equal Correlations," Psychometrika, 16 (March 1951), 3-11.

Mosteller, Frederick. "Remarks on the Method of Paired Comparisons: III. A Test of Significance Assuming Equal Standard Deviations and Equal Correlations," Psychometrika, 16 (June 1951), 207-18.

Neidell, Lester A. "Procedures for Obtaining Similarities Data," Journal of Marketing Research, 9 (August 1972), 335-37.

Rao, Vithala R. and Ralph Katz. "Multidimensional Scaling Methods for Large Stimulus Sets," Journal of Marketing Research, 8 (November 1971), 488-94.

Reif, Rita. "Sheets by Design," New York Times (Sunday, February 13, 1977), F-1 ff.

Reynolds, William H. "Some Empirical Observations on a Ten-Point Poor-to-Excellent Scale," Journal of Marketing Research, 3 (November 1966), 388-90.

Taylor, James W., John J. Joulahan and Alan C. Gabriel, "The Purchase Intention Question in New Product Development: A Field Test," Journal of Marketing, 39 (January 1975), 90-92.

Thurstone, L. L. "Psychophysical Analysis," Journal of American Psychology, 38 (July 1927), 368-89.

UNDERSTANDING UNIVERSITY CHOICE:
A MULTI-ATTRIBUTE APPROACH

Ronald Vaughn, Bradley University
Joseph Pitlik, Bradley University
Behram Hansotia, Bradley University¹

Abstract

This article presents a detailed analysis of the components of a university's image. University choice criteria are studied as to their relative determinance (using a new modified determinant attribute approach). The underlying cognitive structure of university choice is also investigated via factor analysis.

Introduction

The need for the application of marketing theory to college admissions work has become apparent in recent years as the market has changed from the sellers market of the 60's to a buyers market for the 70's. Available information at present seems to indicate that the market for prospective college students will remain a buyers market into the 1990's and perhaps beyond.

Several factors external to the university system have contributed to this shift. The pool of applicants available from high school classes has peaked with the graduation class of 1970 and is expected to decline steadily through 1983. Furthermore, the abolishment of the draft and the ratio of college graduates to job openings in most academic majors appear to be vital forces in the prospective students initial decision as to whether to attend college. In the present competitive environment the four year institution now faces several factors that were almost nonexistent in the 1950's. The early 60's were characterized by the rapid development of community colleges, and closely followed by the development of technical programs offered at two year institutions. These two factors taken together have further reduced the pool of applicants or market potential available. The private university has yet another environmental threat. The large state university, long dormant in aggressively recruiting students, has awakened in recent years and appears to becoming active in its search for students to meet enrollment objectives.

The implication is clear. To successfully compete in the educational marketplace of the present and future, universities, especially private universities, must begin refining their marketing strategies. These refinements will often be in the form of tactical shifts in the promotion and product areas, and a sound base of research is almost a prerequisite to guide these strategies and tactical changes.

Some Problems

Image research has had wide use in marketing generally, e.g., company and brand studies, and in numerous store

imagery studies. In fact, since 1960, more than forty image-related articles have appeared in the marketing literature, more than half of these in the last five years. Further, several environmental trends seem to dictate that this type of research will intensify in the future due to the increasing sophistication of the shopper (including the educational shopper), increased shopper mobility, product proliferation, and other competitive pressures. Just as retailers are facing an intensification of inter-type competition which is causing retailers to combat retail sectors whose merchandise had never before competed with theirs, so to higher education's already declining market is being impinged upon by proprietary and industrial schools. Not only does IBM, General Electric and other corporations now offer bachelor's degrees, but there is increased competitive pressures from community colleges and other more traditional colleges. An article by Kotler and Levy (1969) point out that marketing is a pervasive societal activity and that student recruitment by colleges is even a marketing activity.

Despite the apparent need for imagery research on non-profit educational institutions and despite the wealth of imagery research on the profit-oriented business sector there has been a paucity of research dealing with the marketing problems of a university and hardly any research specifically investigating the image of a university. Although Kotler and Levy (1969) noted much earlier than nonprofit organizations have been typically ignored by the student of marketing, most of the university-oriented marketing literature to-date has been quite general and largely conceptual (e.g., Kotler 1976, O'Brian 1973, Berry and George 1975, etc.), except for a very limited study by MacLachian and Leister (1975) which used multidimensional scaling to develop perceptual maps of the institutional image and a more extensive study of a graduate school's image by Naidu (1969).

This paper attempts to provide insight on the image of a university. Specifically, evidence will be contributed on some of these issues. Can a university's image be measured along salient image dimensions? If so, which attributes are most important? Do determinant attributes exist? How is the information about a center structurally organized, i.e., rather than just describing a university as very good or bad, how is information about a university coded and grouped into cognitive categories? Studying these relatively new issues to educational choice behavior should provide insight as to how educational institutions are evaluated with respect to each other. Finally, understanding imagery development and the choice process in a university setting should serve as a prerequisite for developing better diagnostic and predictive models. While this paper has not been intended as a dissertation on the usage and development of multi-attribute models, a number of reviews have been done which would be quite helpful in later attempts to model the educational choice process, e.g., Cohen (1972), Hansen (1976), and Pessemier and Wilkie (1972). The conceptual article by Kotler (1976) would also be quite helpful in developing more explicit models of educational choice. Towards this objective, a variety of issues and directions for future research are provided in the final section of this paper.

¹Ronald Vaughn, Joseph Pitlik and Behram Hansotia are, respectively: Associate Professor in the Department of Marketing; Instructor in the Department of Marketing; and Associate Professor in the Department of Business Management and Administration. All three professors are at Bradley University.

Methodology

To determine an initial set of criteria that represents the salient components of university preference a preliminary study was conducted. Freshman College of Business Administration students at Bradley University were asked to give the five most important criteria they used in evaluating a university. This questioning process produced an extensive list of criteria that was further modified through discussions with the admissions director and associated admissions personnel plus a review of the literature. A final set of 16 criteria, shown in Table 1 was utilized for further study. While not

ed sample, the results presented are somewhat tentative and additional research is needed.

Results

One of the most complex processes the student and his parents go through is in making the final choice among colleges that have accepted him or her. In this final choice process the student is believed to rate each college's characteristics on the criteria they hold to be most important. Thus, basic to developing proper college marketing procedure is an understanding of the uni-

TABLE 1
UNIVERSITY SELECTION CRITERIA IMPORTANCE

Decision Criteria	Total Sample Importance Mean	Importance Rank	BU Attribute Evaluation Mean	Other Private Universities Attribute Evaluation Mean	State Universities Attribute Evaluation Mean
X ₁ Quality of Education Received	6.449	1	5.821	5.559	5.415
X ₂ Quality of Faculty	6.345*	2	5.678	5.459	5.317
X ₃ Reputation of the Business Program	6.353*	3	5.864	5.041	5.133
X ₄ Academic Reputation of University	6.336*	4	5.782	5.597	5.489
X ₅ Amount of Individual Assistance That Could Be Provided by the Faculty	6.069	5	5.107	5.105	4.000
X ₆ Number and Variety of Courses Offered	5.828	6	5.586	5.125	5.683
X ₇ Size in Terms of Number of Students per Class	5.707	7	5.707	5.451	4.199
X ₈ Size in Terms of the College's Student Population	5.595	8	5.756	4.823	4.044
X ₉ College Displayed a Personal Interest	5.560**	9	5.248	4.869	3.826
X ₁₀ Basic Cost of Attending	5.560**	10	4.364	4.648	5.165
X ₁₁ Location	4.483	11	5.424	4.853	4.956
X ₁₂ Job Placement Services (after graduation)	5.422	12	5.379	4.660	4.692
X ₁₃ Campus Size	5.353	13	5.639	4.702	4.207
X ₁₄ Housing Facilities	5.000	14	5.060	4.669	4.610
X ₁₅ Availability of Financial Aid	4.491	15	4.734	4.702	4.369
X ₁₆ Influence of Friends Who Are Attending	4.078	16	4.153	4.195	4.177

*These three variables tied for the second rank.

**These two variables tied for the ninth rank.

exhaustive, these criteria are thought to represent those criteria most commonly used by prospective students and their parents. A questionnaire was then designed to determine the relative importance of these criteria, to both parents and students. Respondents were asked to rate in importance each criteria on a seven point scale where a "7" was extremely important. Respondents were also asked to identify and then evaluate each of their top three choices along the same criteria on a scale from 1-7 where a "1" was a poor evaluation and a "7" was a good evaluation.

In January of 1975 402 questionnaires were mailed to all students and parents of students who were admitted to the College of Business Administration for the Fall of 1974. The 30% response rate is primarily attributed to the rather low (16%) response rate in the category of "Parents of Students Admitted, but not Enrolled." However, the factors of a lengthy questionnaire and no monetary incentive provided for the respondent also probably influenced the low response rate. Due to the limit-

versity choice process, which, in turn, must be grounded in a thorough analysis of what attributes of a college are evaluated. We will attempt to provide insight to the university evaluation process by several analysis stages. The first stage of analysis will determine the relative importance of sixteen university choice criteria. Second, the evaluation of several alternative universities on each attribute will be compared. Third, selection criteria will be evaluated for relative determinancy, and fourth, the underlying cognitive structure of university choice will be explored using factor analysis.

Relative Importance of University Selection Criteria

Sixteen choice variables were selected for importance analysis. Table 1 shows that the variable means ranged from a low of 4.060 for the influence of friends who are attending the school, to a high of 6.449 for the quality of education received. Furthermore, the typical respondent put a premium on the quality of faculty, repu-

tation of the business programs, academic reputation of the university, amount of individual assistance that could be provided by the faculty, number and variety of courses offered, the number of students per class and so on. These findings appear somewhat consistent with those of Holland (1958 and 1959) in his studies of what aspects of a college made it "best." Using a sample of parents of National Merit Scholars they found the key aspects to be quality of faculty, scholastic standards, curriculum, reputation and facilities. However, in the actual choice of college "practical" and "financial" factors were of substantial influence with academic factors playing a moderate, secondary role. Knowledge as to which attributes are most important undoubtedly suggests that certain educational attributes must be particularly considered for attribute development, maintenance and promotional activities.

Institutional Evaluation

The marketing task of a college is to articulate a mission and purpose for the college that makes sense in view of its history, resources, opportunities, competition and other situational variables (Kotler, 1976). Further, a coordinated marketing effort must assist in developing a distinct posture for the university, i.e., an institutional position that will meet the educational needs and desires of its particular market segment as a means to satisfying institutional goals. With this in mind, one of the first steps in developing an institutional position is to measure and evaluate the university's current perceived position in the market in relation to competitors. While knowledge of which attributes are most important provides some direction for faculty and administrative efforts at developing and maintaining its position, periodic assessment of the university's current perceived position provides additional insight to institutional strengths and weaknesses and suggests areas of the educational product where improvement or change might be necessary. As Gould (1971) points out, constructive change by a college must be preceded by proper introspection: "higher education must be more truthful in its own evaluation of its current weaknesses in order to portray itself effectively to its many constituencies."

Stage two of the analysis was to evaluate Bradley University's image on each of sixteen criteria and compare it to the perceived image of several competing colleges (grouped into an other private and a state school classification). While it was expected due to the nature of the sample (people who had applied and been admitted to Bradley University), Bradley had a more positive evaluation than its competitors on most of the attributes. It should be noted that other publics of the university if sampled might have provided different average evaluations. Table 1 shows some attributes on which Bradley seems to be viewed as quite distinct while on others there seems to be an insignificant perceived difference. Bradley seems to be viewed as relatively different from its competitors (in a positive sense) on size of the student population, class size, college displayed a personal interest, the amount of individual assistance provided by the faculty, campus size, housing facilities and reputation of the business program. Bradley was viewed as quite different from its competitors in a negative sense on only the basic cost of attending. Only a slight variation among the evaluation seemed to exist on the influence of friends attending, academic reputation of the university, quality of the faculty, quality of the education received and availability of financial aid.

While this study of perceived evaluations by attributes is helpful in pinpointing relative strengths and weaknesses, additional information is needed. Although a university is perceived as quite different in a positive

sense from its competitors (on a given attribute), a relative strength does not necessarily exist unless that attribute happens to be of considerable importance to the educational consumer. Likewise, the fact that a university is viewed more poorly is not necessarily cause for alarm unless the criterion happens to be of crucial importance. Thus, this analysis suggests that a university should first, discover the relative importance of the various university selection criteria; second, investigate the image of themselves and their competitors on these same criteria; and, third, analyze the importance and evaluation scores in combination. The analysis of determinant attributes presented in the next section illustrates a way of analyzing attribute importance and evaluation together.

Determinant Attributes

The third stage of the analysis was to assess the principal university decision variables and their relative determinance in university selection decisions. Determinant attribute analysis should help to understand those features which move the educational consumer to action, i.e., to make him prefer a particular university, actually attend there and/or recommend it to friends. Because of their crucial relationship to preference, a university's perceived position with respect to these determinant attributes must be closely investigated to check for strengths, weaknesses, opportunities, etc.

Attribute determinance was estimated by evaluating (1) the importance associated with a particular university selection criterion and (2) the degree to which competing universities were perceived to differ in terms of that selection criterion (Myers and Alpert, 1978). While these two aspects of a determinant criterion are part of the original definition of "determinance," the operationalization of this concept represents a new approach. Using Hansen's (1977) "modified determinant attribute analysis" each individual's response on a given attribute was cross-tabulated based on their response to each of the two questions (i.e., importance and perceived variation). Table 2 presents a sample of such a cross-classification for a product attribute, where the table cell values are reported as percentages. Hansen (1977) presents a critical evaluation of this new operationalization compared to the more traditional operationalization of attribute determinance. Using the Myers and Alpert definition of determinance, the attribute having the highest proportion of respondents classified in cell 9 of Table 2 where both importance and variation are high, would be the currently most determinant analysis was based on the present or current perceptions of the product. Cell 7 of Table 2 is also of interest as it presents the proportion of the respondents who indicated the attribute was high in importance, but low in perceived inter-university variation. This cell 7 is referred to as "potential" determinance because if one of the competing universities could produce an educational product that is perceived or evaluated much higher than its competitors, then that university would have a distinct advantage. This is because the current high importance-low variation attributes may offer the greatest potential for improvement since an increase in a university's evaluation on that dimension would lead to a distinct perceived difference on a "potentially" crucial or determinant choice criteria. The other cell values in Table 2, e.g., 6, 5, etc. have similar implications in that strategies may be suggested to increase the university's evaluation, or to increase the perceived importance of an attribute, or both. The remainder of the analysis, however, concentrates on cell 7 and cell 9.

The results of the determinant attribute analysis are shown in Table 3. Several attributes stand out as

TABLE 2
SAMPLE CROSS-CLASSIFICATION
(Cell numbers are only for identification)

		Perceived Variation of Attribute		
		Low (0-1.000) ^d	Medium (1.001-2.000) ^e	High (2.001-or greater) ^f
Importance of Attribute	Low (1-4) ^a	1	2	3
	Medium (5) ^b	4	5	6
	High (6-7) ^c	7	8	9

^a"Low" importance represents a 1-4 scale value (extremely unimportant, somewhat unimportant and neither important or unimportant).

^b"Medium" importance represents a 5 scale value (somewhat important).

^c"High" importance represents a 6 or 7 scale value (very important or extremely important)

^d"Low" variation represents a standard deviation across alternative universities of one or less scale value standard deviations.

^e"Medium" variation represents a standard deviation across alternative universities of 1.001 to 2.000 scale value standard deviations.

^f"High" variation represents a standard deviation across alternative universities of greater than two scale value standard deviations.

currently relatively determinant: college displayed a personal interest (22.81%); size in terms of the college's student population (20.86%); campus size (19.83%); and, amount of individual assistance that could be provided by the faculty (19.30%). The potentially determinant university choice criteria include: academic reputation of the university (63.94%); quality of the faculty (63.82%); reputation of the business programs (55.46%); and, the number and variety of courses offered (46.22%). It must be noted, however, that the variables that came out as currently or potentially determinant in this sample are quite possibly not determinant or crucial to the selection process of other universities. Remember that the sample being studied in this case was students who had applied to and were admitted to the College of Business Administration at Bradley University (and the majority of these students actually enrolled). Thus, the nature of the sample suggests that the currently determinant criteria are more likely to be those variables that were crucial to the decision to attend Bradley. The fact that Bradley is evaluated much higher than its competitors in Table 1 lends additional credibility to this interpretation. On the other hand, the potential determinant criteria suggest attributes that could contribute more to Bradley being chosen if the perceived evaluation of Bradley could be improved to the point where it was viewed as quite different from its competitors. Table 1 also lends credibility to this interpretation as the evaluations of Bradley University on the potentially determinant criteria show relatively little difference from the evaluations of its competitors.

TABLE 3
RESULTS OF CROSS-TABULATION
CALCULATION OF ATTRIBUTE DETERMINANCE

Attribute	Proportion in High Importance High Variation Cell	Proportion in High Importance Low Variation Cell
Personal Interest	22.81%	26.27%
Size of College's Student Population	20.86%	21.01%
Campus Size	19.83%	15.25%
Amount of Individual Assistance by Faculty	19.30%	36.97%
Location	16.10%	24.37%
Number of Students per Class	15.65%	32.20%
Reputation of Business Programs Job Placement Service	13.91%	22.03%
Housing Facilities	11.01%	23.53%
Quality of Education Received	10.62%	23.73%
Academic Reputation of University	8.85%	63.94%
Quality of Faculty	8.77%	63.82%
Basic Cost of Attending	8.54%	33.05%
Number and Variety of Courses	6.96%	46.22%
Availability of Financial Aid	6.78%	26.27%
Influence of Friends Attending	4.27%	9.40%
Cognitive Structure of the University		

In stage four of the analysis a factoring procedure was used to determine the underlying constructs or dimensions which the respondents considered when selecting a university. The same 16 selection criteria listed in Table 1 were the subject of the factor analysis. PA2, the iterative procedure for determining communalities, was used from the SPSS package. In 23 iterations, five factors with eigenvalues greater than one were extracted. The fifth factor explained only 6.4% of the total variance and since after the varimax rotation the fifth factor had no significantly high factor loadings on it, we decided to drop the fifth factor. An "L" test as described in Van de Geer (1971) also confirmed our dropping the fifth factor, which had an eigenvalue of only 1.016.

The first four factors explained 56.3% of the total variance. Reported below is the matrix of factor loadings after the varimax rotation. The asterisks (*) adjacent to the higher factor loadings are placed to help identify the important attributes of each salient university selection dimension. Factor 1, for example, is loaded heavily on variables 2, 5, 3, 4, and 1, which are: quality of faculty, amount of individual assistance provided by faculty, reputation of the College of Business Program, academic reputation of the

TABLE 4
MATRIX OF FACTOR LOADINGS

Attributes	Factor 1	Factor 2	Factor 3	Factor 4
X ₁	0.652*	0.176	0.029	0.040
X ₂	0.654*	0.137	0.069	-0.158
X ₃	0.677*	-0.012	0.174	0.163
X ₄	0.604*	0.048	0.138	0.226
X ₅	0.562*	0.168	0.179	-0.146
X ₆	0.490	0.015	0.220	0.325
X ₇	0.235	0.547*	0.101	-0.168
X ₈	0.025	0.695*	1.019	0.190
X ₁₀	0.016	0.255	0.878*	-0.039
X ₁₁	1.107	0.164	-0.016	0.622*
X ₁₂	0.252	-0.133	0.337	0.010
X ₁₃	-0.071	0.847*	0.010	0.220
X ₁₄	0.108	0.442	0.092	0.004
X ₁₅	0.166	0.026	0.587*	0.098
X ₁₆	0.145	0.270	0.266	-0.016

university, and quality of education received. Apparently the first dimension is "university and College of Business Program quality." Interestingly, the respondents do not seem to look at only the quality of the Business program but also at the quality of the entire university. Furthermore, they seem to relate the amount of individual assistance to quality.

The second factor is loaded heaviest on variables 8, 13, and 7; which are size in terms of the college's student population, campus size, and size in terms of the number of students per class. Obviously, this is a "size" dimension.

The third factor is loaded on variables 10 and 15, which are basic cost of attending and availability of financial aid. This seems quite clearly to be a "total university cost" dimension. Thus, the respondents do not seem to look at just the cost of attending a university, but at the total cost obtained by offsetting tuition and housing costs by availability of financial aid.

The fourth and last factor has a high loading on only the eleventh variable, namely location. Hence, it can be labeled a "location" dimension. Note that of the 16 variables originally selected for the analysis (see Table 4), 5 did not have appreciably high loadings on any of the four factors. They are variables 6, 9, 12, 14, and 16. Of these, variables 12 and 16, namely job placement service after graduation and influence of friends who are attending, are least correlated to any of the four factors discussed above. Variables 6, 14 and 9 on the other hand do have modest correlations with factors 1, 2 and 3 respectively. Implying that variable 6 which is number and variety of courses offered is moderately correlated with the "quality" dimension. Variable 14, which is housing facilities, is moderately correlated with the "size" dimension and variable 9, college displayed a personal interest, is correlated with the "total cost" dimension.

Table 5 summarizes our findings in this section.

Research Implications

To continue to improve our understanding of the university choice process, research is needed on many issues, including the following:

1. How do university choice criteria differ between the parent (an important decision influencer) and the student? Are there differences in attribute saliency or determinancy between the student and the parent? Further, are there such differences for students who apply and do not enroll versus those who do enroll, i.e., users versus nonusers.

TABLE 5
SUMMARY OF FACTOR ANALYSIS

Four factors accounted for 56.3% of the total variance:

1. "University and Business Program Quality" Index, described by:
 - a. quality of faculty
 - b. amount of individual assistance
 - c. reputation of the Business program
 - d. academic reputation of the university
 - e. quality of education received
2. "Size" Index, described by:
 - a. size in terms of student population
 - b. campus size
 - c. size in terms of the number of students per class
3. "Total University Cost" Index, described by:
 - a. basic cost of attending
 - b. availability of financial aid
4. "Location" Index, described by location

Also, are there differences in saliency or determinancy for students of differing quality, i.e., for students with very promising academic potential versus average academic potential (Naidu suggested in 1969 that due to a greater number of alternatives, financial aid is more important to attracting students of high quality)?

2. What image does the university have among its other important publics, e.g., alumni, donors, high school and junior college counselors, and the community at large? As Shapiro (1972) suggests an institution such as a college is concerned with two primary sets of transactions or exchanges, one is with donors and the other is with clients and for a college highly dependent on external funding sources the donor group may be at least as important as the client group (students) to study and understand.
3. Do university selection criteria change (i.e., are there new criteria) after the prospective student enrolls and attends the college for some time? Perhaps more subtly do criteria saliences or evaluations change through time? A study by Naidu (1969) suggests that the "attractiveness" of graduate education varies from level to level, e.g., he notes that financial aid and faculty reputation are considered to be more important by the student at graduate levels, but at the undergraduate levels location of the campus and general reputation of the university seemed more important in evaluating the educational product. It may be particularly important to understand the differential needs of seniors so that they will graduate as satisfied customers. Also, a more thorough analysis of the needs or disconfirmed expectancies of sophomores and juniors may help reduce the problem of retaining students for four years.
4. Are the perceptions of university faculty, administrators and managers consistent with the perceptions of their publics? As the images held by collegiate personnel help to determine the university's projected image to other publics, this aspect of image may also be of real concern.
5. What are the key information sources which students and other publics use, and which are the basis for attitude formulation?

6. Can the educational market be meaningfully segmented? Just as in other businesses, a school needs to profile its perhaps rather small potential market as a guide to effective and efficient allocation of resources.
7. Once a university has assessed its current perceived position, what strategies and tactics should it consider? Kotler (1976) suggests that such strategies and tactics go beyond just the work performed by the admissions office, but should include such activities as institutional positioning, portfolio planning, applicant development work, applicant evaluation and notification work, recruitment effort evaluation, college improvement planning, and alumni loyalty work. Little empirical work has been done to determine the effectiveness of different strategies or ways of implementing them. Likewise, the differential sensitivity of different market segments to certain marketing strategies or tactics has had little focus by marketing people.

In summary, while there has been a considerable amount of research focused on nonprofit organizations by marketing people in the past 10 years (as illustrated by Rothschild's (1976) 43-page bibliography of works relating to marketing for public sector and nonprofit organizations), there has been a lack of focus on the educational sector, which is one of the largest and most important service industries in the U.S. It is hoped that the reported research study and suggestions, along with consideration of the benefits that may accrue directly to academicians, will prompt a more active interest in this area by marketing researchers.

REFERENCES

1. Berry, Leonard L. and William R. George. "Marketing the University: Opportunity in an Era of Crisis," Atlanta Economic Review, 25 (4) (July-August 1975), 4-8.
2. Cohen, Joel B. "Toward an Integrated Use of Expectancy-Value Attitude Models," Paper presented at the ACR/AMA Workshop on Buyer/Consumer Information Processing, University of Chicago, 1972.
3. Gould, Samuel B. "Bridging the Interpretation Gap," College and University Journal, 10 (May 1971), 14-16.
4. Hansen, Flemming. "Psychological Theories of Consumer Choice," Journal of Consumer Research, 3 (December 1976), 117-142.
5. Hansen, Robert A. "Determinant Attribute Analysis and the Search for Product Modification Ideas: Theoretical and Managerial Considerations," an unpublished working paper, Ohio State University, 1977.
6. Hie, Norman, et al. Statistical Package for the Social Sciences (SPSS). New York: McGraw-Hill, Inc., 1975.
7. Holland, John L. "Parental Expectations and Attitudes about College," College and University, 34 (Winter 1959), 70.
8. Holland, John L. "Student Explanations of College Choice and Their Relation to College Popularity, College Productivity, and Sex Differences," College and University, 33 (Spring 1958), 313-20.
9. Kotler, Philip and Sidney J. Levy. "Broadening the Concept of Marketing," Journal of Marketing, 33 (January 1969), 10-15.
10. Kotler, Philip. "Marketing Theory and Its Application in College Admissions," paper presented at the Colloquium on College Admissions, The Abbey, Lake Geneva, Wisconsin, May 1976.
11. MacLachlan, Douglas L. and Douglas V. Leister. "Institutional Positioning: A Metamarketing Application," Marketing's Contribution to the Firm and to the Society, ed. R. C. Curham. Chicago, American Marketing Association, 1975.
12. Myers, James H. and Mark I. Alpert. "Determinant Buying Attitudes: Meaning and Measurement," Journal of Marketing, 32 (October 1968), 13-20.
13. Naidu, Gurramkonda M. Systems Approach to the Marketing Aspects of Higher Education, an unpublished dissertation at Michigan State University, 1969.
14. O'Brian, Edward J. "Marketing Higher Education," College and University Journal, 12 (4) (1973), 22-23.
15. Pessemier, Edgar A. and William L. Wilkie. "Multi-Attribute Choice Theory--A Review and Analysis," Institute Paper No. 372, Institute for Research in the Behavioral, Economic and Management Sciences, Purdue University, 1972.
16. Rothschild, Michael L. "An Incomplete Bibliography of Works Relating to Marketing for Public Sector and Nonprofit Organizations," an unpublished working paper, University of Wisconsin--Madison, January 1976.
17. Shapiro, Benson P. "Marketing in Nonprofit Organizations," an unpublished working paper, Marketing Science Institute, 1972.
18. Van de Geer, John P. Introduction to Multivariate Analysis for the Social Sciences. San Francisco: W. H. Freeman and Company, 1971.

AN EXPLORATORY STUDY OF THE INFLUENCES OF USAGE
SITUATION ON PERCEPTIONS OF PRODUCT-MARKETS

Rajendra K. Srivastava (student), University of Pittsburgh
Allan D. Shocker, University of Pittsburgh
George S. Day, University of Toronto

Abstract

Recent research has illustrated the importance of environmental influences in determining consumer behavior. A typology of relevant "objective" usage-situations is developed, and subsequently, the composition of competitive product markets determined based on the similarity of the usage patterns. Results appear to show the feasibility of creating meaningful product specific situational taxonomies.

Introduction

Consumer research in marketing, heavily influenced by a psychological tradition, has typically followed a paradigm where consumer choice in the market has been examined largely as a function of personal characteristics and preferences. In recent years, both the psychological tradition (e.g., Fredriksen (1972), Wicker (1972), Bowers (1973), and Endler and Magnusson (1976)) and consumer behavior research, (Belk (1975), Lutz and Kakkar (1975), Bearden and Woodside (1976), and Hansen (1976)) have examined the evidence for situational effects moderating consumer behavior and have proposed conceptual models of these effects. This work has generally supported the hypothesis of important situational effects but raised problems of measurement and classification.

Theoretical development in the area of situational influence on consumer behavior is still in an early stage. Although there appears little doubt that environmental factors affect behavior (see reviews by Lutz and Kakkar (1975), Belk (1975), and Endler and Magnusson (1976)), there exist differing viewpoints as to the line of inquiry which would best illustrate the interaction between individual characteristics and situations. Perhaps one of the more limiting drawbacks at the current stage of development is the lack of consensus on a classification schema or typology for situations within which consumer behavior may be systematically studied (Wicker (1972), Fredriksen (1972), Belk (1975), Russell and Mehrabian (1976)). Without such a comprehensive typology, systematic analysis of the influence of situations on choice behavior cannot be achieved. The two primary approaches to the development of such typologies have not been totally satisfactory. The first has sought to analyze objective situations (those meaningful to respondents) to find situational factors which appear to affect behavior (e.g., Belk (1975) proposed five groups of objective characteristics-- physical surroundings, social surroundings, temporal perspective, task definition, and antecedent states). The second has sought to classify situations by the nature of the psychological processes to which they give rise (e.g., Lutz and Kakkar (1975) attempted to classify situations in terms of Mehrabian and Russell's dimensions of pleasure, arousal, and dominance). Allen (1965), Barker (1975) and Wicker (1975), among others, have discussed variants of these approaches.

Typologies which retain an objective character have the advantage of greater meaningfulness (i.e., they provide some usable explanations for situational effects on behavior). Its major disadvantage is that the potentially large number of dimensions may lead to a correspondingly large number of distinct situational types.

The psychological situation in turn offers parsimony at the possible expense of relevance. However, for most managerial marketing purposes, one might expect that the psychological situation would prove less usable (Lutz, personal communication).

The focus of this paper is on the usage or application situation; that is, the objective circumstances for which the product or service is purchased. Specifically, the objective is to develop a situational typology that can account for a comprehensive array of usage situations. In other words, we are seeking a method for extracting the minimum number of underlying dimensions which account for the variability in the effects of a large number of usage situations. Not only is such a typology useful in clarifying the general concept of usage situation, it can also be used to suggest additional situations to incorporate in the analysis.

In discussing alternative criteria for the development of a classification of situations (centered upon various taxonomic procedures such as cluster and factor analysis), Fredriksen (1972) suggested that one could be based upon elicited behavior. The consideration of products or brands for use under various situational contexts would appear to be a worthwhile organizing criterion for such a classification. In this study the setting is consumer perceptions of competitive product-markets.

The composition of competitive product-markets is of considerable practical relevance for a number of managerial and public policy questions (Day and Shocker (1976)). Several researchers have argued for definitions of markets based on consumer behavior or judgment (e.g., Jain and Etgar (1975); Bourgeois, Haines and Sommers (1975)). By and large, these procedures are based upon measures of overall perceived product similarity and have not explicitly considered the effect of usage situations. Since competitive products are more or less appropriate for the same usage situations, a refinement of the approach to include situational effects would be desirable. Thus a corollary purpose of this paper is to develop methodology to implement this concept.

Though this is an exploratory study, there is some precedent for this general line of thought in the work of Steffire (1971) who found analysis of product-usage associations beneficial in operationally defining product-markets for the purpose of new product development. His purpose was not to create a situational typology but rather to find groupings of products which were similar in terms of their associations with a common set of uses. Nonetheless, his efforts would appear to provide some legitimization for our present efforts which can be considered a refinement and extension of his. As justification for further research lies in the validity of our findings, care has been taken at various stages to assess this. Of course, validation efforts need replication over a broader array of product-market contexts and over more diverse market segments.

Methodology

The Methodology progressed through two stages: (1) Determination of varieties of possibly related products and usage-situations and development of a products-by-usage situations matrix (such that each cell (ij) of the matrix provides a measure of the degree of appropriateness of product i for situation j). Subsequently, taxonomic procedures are used to examine the similarities between pairs of situations (in terms of products judged appropriate), to produce an initial situational taxonomy (and definitions of the competitive product-markets). (2) Refinement of the above taxonomy by development of morphological situational descriptions, testing whether these descriptions were interpreted by a new set of respondents in a manner similar to the researchers' conceptualization, and finally replicating the procedure in the previous step with the more complete morphological usage-situational descriptions.

The "breath freshener" market was chosen for the purpose of this exploratory study because: it (1) showed promise in its ability to show variability in product choice across usage situations, and (2) would be a product class familiar to a convenience sample of socially-conscious graduate students in business administration. The latter consideration is especially relevant as familiarity/experience with the product class would be important in the development of meaningful, realistic usage-situations and for minimizing "guesswork" on the part of respondents in indicating appropriateness of products for specific usage-situations (thus reducing the error component of individual differences between subjects).

Stage 1

A procedure adapted from Stefflre (1971) was used to determine the initial set of products and usage situations. Ten graduate students in administration at a major Eastern university were asked by questionnaire to go through an iterative procedure: from this starting point of a "core" of several breath freshening products and examples of usage situations, they were asked to identify additional situations where use of the products would be relevant and other products appropriate to the expanded set of situations, etc. This led not only to lists of commercial products/brands, but also numbers of home remedies (e.g., rinsing mouth with saltwater, sucking a lemon). An alternative data generation method involved group interviews with an additional 8-10 students which, while leading to additional products and situational descriptions, also explored the choice criteria used in determining appropriateness. Generally, the interviews provided a richer data source than had the questionnaires.

Our own judgments supplemented the insights gained from the interview and questionnaire. Duplications and redundancies in products and situations were eliminated. Additional products were added by surveying the product racks of supermarkets and drug stores. In identifying products, an attempt was made to consider all distinctions between brands that might prove relevant (e.g., breath sprays were distinguished from drops, brands such as Close-up--a brand with the flavor and clarity of a mouthwash--were explicitly identified, different flavors were distinguished). Further, a few "extra" products were included in the list to provide some validation for subsequent procedures (the expectation being that such products would be judged inappropriate). This process yielded a set of 18 situations and 46 products.

EXHIBIT 1

Examples of Products .

"Seven-Up" ("extra" product)	"Listerine" mouthwash
Fruit flavored chewing gums	"Dynamints"
"Roloids" (extra" product)	"Close-Up" toothpaste
"Ultrabrite" toothpaste	"Water-Pik"
"Binaca" breath spray	Mint flavored chewing gums

EXHIBIT 2

Examples of Situations, Stage 1

- To clean mouth upon arising in the morning.
- Before an important business meeting late in the afternoon.
- Before going to a formal party.
- When you have a sore throat.
- To get rid of a stale taste - anytime.
- For use during a social event - may have to offer to acquaintances.
- Before an unexpected business meeting with a client.

An additional sample of 46 graduate student respondents was then asked for each of the 18 usage situations to indicate whether each product was (a) appropriate, (b) inappropriate, or (c) no opinion. Given the large number of judgments required (18 x 46 = 828) and to relieve tedium, this was accomplished by presenting respondents with 46 cards (with product names and identification numbers). After shuffling the cards (once for each situation) and sorting them (products) into the above categories, this information was coded on a questionnaire form.

A Measure of Appropriateness of Product i for Situation j. By aggregating responses across subjects, it was possible to obtain a products-by-usages matrix such that within each product-situational cell we would have counts for the number of respondents judging (1) appropriate (A_{ij}), (2) inappropriate (I_{ij}), and (3) no opinion (N_{ij}). Define:

$$a_{ij} = \frac{A_{ij}}{A_{ij} + I_{ij}}$$

as a measure of the degree of appropriateness of product i for situation j. This measure excludes the "no opinion" responses (which were very infrequent) and merely reflects the proportion of people who considered a product appropriate for a particular situation.

The Test for Perceptual Homogeneity. If there were complete agreement among the subjects, virtually all (or none) would rate the product as a appropriate. Let:

$$v_{ij} = |a_{ij} - 0.5| \quad \begin{matrix} (= 0 \text{ if complete disagreement}) \\ (= 0.5 \text{ if complete agreement}) \end{matrix}$$

be a measure of the level of agreement between respondents for cell ij of the product-use matrix. By estimating the mean and variance of this measure (across all cells), one can compute whether the level of agreement (homogeneity) is "adequate" (for example, significantly different from 80% or 90%, *i.e.*, $v_{ij} = 0.40$ or 0.45 respectively). For Stage 1, v_{ij} has a mean of 0.31 and a variance of 0.0208 . This allows us to reject at the $\alpha = .01$ level the hypothesis that the entries in each cell occur purely by chance. At the same time, we cannot reject the hypothesis of nearly complete agreement (say, $v_{ij} = 0.45$).

Taxonomic Approach. In order to develop a taxonomy of usage-situational influences, we analyse the similarity between the situations (in terms of the products/brands associated with them). One such measure of similarity would be the correlation between columns (situations) in the products-by-usages matrix. This correlation matrix could be used as input into a variety of dimension reducing (taxonomic) techniques. Principal components analysis was the method adopted for the following reasons:

- (1) it was felt that since the input measures (degree of appropriateness) were ratio-scaled, a metric scaling technique was applicable.
- (2) it allowed us to examine the interrelationships between situations and products in the same reduced space (see below).
- (3) because some correlations between situations were very high, the use of a technique such as nonmetric multidimensional scaling would be problematic due to the associated problem of degeneracy.
- (4) further, scaling techniques are primary informative about large distance similarities (each point in a scaling configuration is located by reference to all other points, so that large distance similarities dominate) while clustering techniques are more informative about small distance relationships (Kruksal (1972)).

Since for the purpose of this exploratory study, we are more interested in the interrelationships between generic product types and the association between a set of usage situations, it would be more beneficial to use principal components analysis (metric scaling) to examine the large distance similarities. Subsequently, situational clusters could be formed (based on patterns of factor loadings) and product clusters generated (based on the pattern of factor scores). In addition to the development of a taxonomy of the objective usage-situational influences, these situational clusters would serve as a reasonable basis for product-market definition *i.e.*, product clusters closely associated with the same situational cluster being viewed as competitive and constituting a relevant market.

Analysis (Stage 1). Principal components analysis was conducted with the (18) usage situations as variables and the (46) products as cases using the BMD08M algorithm. The raw data were the appropriateness scores (a_{ij} 's of products for situations). Varimax rotation was performed to yield a rotated factor matrix that approximated simple structure, and unstandardized factor scores obtained for each of the 46 products. This analysis suggested two or three factors (accounting for 52%, 35%, and 6% of the variance, respectively, in the unrotated factor structure). Interpretation of the rotated factor was as follows:

Component 1 - Social versus Personal Concern (*i.e.*, did the situation involve others or just self-e.g., stale taste in mouth)

Component 2 - Away versus at home (implying privacy and availability of alternatives)

Only one situation ("when you have a sore throat") had high factor loadings on the third component. Since this situation had very low loadings on the first two dimensions it seemed to have little in common with the other situations. Examination of the product-use matrix revealed that products that were generally inappropriate for the other situations (e.g., cough syrups, sodas, and the "extra" products) had comparatively higher measures of appropriateness for this situation. It thus seems that this situation was picking up the variance for the products that were in general considered inappropriate for breath freshening. Accordingly, only two factors were retained.

The unstandardized factor scores for these 46 products were plotted. While not shown here, the results were highly intuitive with toothpastes and mouthwashes having high scores on component 2 (home) and low scores on component 1 (personal concerns). Similarly, candies, mints and gums loaded low on component 2 (away from home) and high on component 1 (social concerns). The "extra" products were not strongly associated with any of the situations thereby indicating their general inappropriateness as breath fresheners.

Discussion. The results of principal components analysis were to a large extent congruent with prior expectations. Commercial products as opposed to "home remedies" were generally associated with breath freshening. Toothpastes and mouthwashes were closely related in terms of their general association with virtually all situations. Breath sprays and drops appeared to be a distinct category. However, there were some things that were not anticipated. An interesting finding was that for several situations (generally centering around social occasions) mint-flavored gum and candy products were seen as more appropriate breath fresheners than fruit-flavored gums and candies. Flavor appeared to be a better indication of product-market differentiation than the nominal product types (candies or gums).

The purpose of this initial stage was to develop a crude taxonomy of situations (and products). The ad hoc elicited situations had been of varying levels of specificity (*e.g.*, to get rid of stale taste in the mouth - anytime) and completeness, and some more clearly called for breath freshening than others. For example, the situation "when you have a sore throat" was not associated to an appreciable extent with breath freshening products. In addition, four situations resulted in possible confusion regarding their interpretation (as indicated by a low level of agreement (v_{ij}) across respondents). In view of these shortcomings, it appeared useful to create a morphological set of usage situations of more uniform levels of specificity and completeness, and then to replicate the procedures described above.

Stage 2

A debriefing session was held with the 46 respondents who had participated at stage 1. They confirmed the ambiguity of several of the troublesome situations, suggested that current descriptions of situations were deficient in not disclosing the amount of time available to freshen breath, and indicated some lack of identification with certain situations. These comments called for changes in the procedures.

Interpretation of the two-factor solution at Stage 1 provided the major guidance for the development of a situational taxonomy. The comments of the students at the debriefing together with our own judgment suggested

two additional situational dimensions. Four dimensions were posited:

1. Personal vs. Social Concerns: "Self" - the primary concern is with the subjects' personal reactions to either their own mouth freshness or to the freshness of others' breath; "Other" - the subjects' primary concern is with others' reaction to their (the subjects') breath freshness. (2 levels)
2. Privacy of Use and Access to Alternatives: "Home" - usage of a breath freshener will be in a location affording considerable privacy of use and ability to use a wide variety of breath freshening alternatives; "Away" - usage of a breath freshener will be in a location affording less privacy of use and more limited ability to use a variety of alternatives. (2 levels)
3. Risk of Being Noticed: "High Noticeability" - a high expectation that the subjects' state of breath will be noticed by others or by self; "Lower Noticeability" - a lower expectation that the subjects' state of breath freshness will be noticed by others. (2 levels)
4. Amount of Time to Prepare: "Time to Prepare" - the subject has adequate time to freshen breath; "Limited Time to Prepare" - the subject has limited time to freshen breath. (2 levels)

The first two dimensions of the typology were obtained from the interpretation of the Stage 1 factor structure. The third and fourth were added because of their apparent plausibility.

New situational descriptions were generated to correspond to all possible (morphological) combinations of the levels of the four characteristics. 27 such descriptions were developed providing duplication for several of the possible (2x2x2x2=16) cases. In order to determine whether these descriptions validly represented their respective cells, successive groups of six to eight students were given each description together with the typology above and asked to code each description back into the typology. After each iteration an attempt was made to understand those instances where there was substantial disagreement among the subjects. Some situational descriptions needed to be rewritten. Two iterations of this process produced substantial convergence. All but three situational descriptions produced a high level of consensus and these three were dropped from further analysis. Unfortunately, some of the changes resulted in increased representation for several cells of the typology and no representation in a few others. In all further analysis, situational descriptions are assigned to the cells of the typology using this "back coding" of the judges. Exhibit 3 shows examples of situations in stage 2.

38 respondents in the Graduate School of Management at a major West Coast University completed the task of indicating appropriateness (appropriate/inappropriate/undecided/product unfamiliar) for the 48 products¹ for 14-15 randomly assigned (from the list of 24) situations. In addition data were collected on the overall effectiveness of each product as a breath freshener and the importance, frequency of occurrence, and "realism" of the situation.

¹Two products were added to the set of products at the first stage.

EXHIBIT 3

Examples of Situations, Stage 2

It is morning and you are preparing to go to work. You expect to be tied up in routine meetings most of the day.

You are preparing to go to a formal dinner at the home of an important client.

You are preparing to go to work knowing you have an appointment with your family dentist later that morning.

You are getting ready to go out on a blind date where you expect intimacy.

You are at home waiting for that important person in your life to arrive for some intimate pleasure.

You are at home preparing to go out to a casual social function with friends.

Measure of Appropriateness. A measure of the degree of appropriateness similar to the one in Stage 1 was designed to account for the inclusion of the undecided category. If A_{ij} = count for appropriate, I_{ij} = count for inappropriate, and U_{ij} = count for undecided, for each cell ij , then:

$$a_{ij} = \left(\frac{A_{ij} + 0.5 U_{ij}}{A_{ij} + I_{ij} + U_{ij}} \right)$$

The counts for U_{ij} (undecided) for each cell were very low compared to A_{ij} (appropriate).

Test for Perceptual Homogeneity. v_{ij} , the measure of the level of agreement, was again calculated. It had a mean of 0.258 and a variance of 0.019. This allowed us again to reject (at the $\alpha = 0.05$ level) the hypothesis that the entries in each cell were due purely to chance. The average level of agreement is lower in this stage. However we must keep in mind that each of the 38 respondents indicated product appropriateness for only 14-15 randomly selected (from the list of 24) situations. Thus on an average, the appropriateness measure (a_{ij}) and consequently the level of agreement (v_{ij}) were calculated on the basis of 25 responses. Consequently, a somewhat lower level of agreement may be expected as each "aberrant" response will have a larger effect on the measure of appropriateness.

Analysis. Again principal component analysis using the situations as variables was conducted. The first three unrotated components accounted for 57%, 30%, and 6% of the variance, respectively. Varimax rotation was again performed to improve interpretability. Examination of the rotated factor structure led to an interpretation of the first two factors as at Stage 1. (The third factor was more ambiguous with only two situations loading greater than 0.4 on it and one of those loaded quite heavily on factor 1). From this we conclude that the two additional dimensions in the typology added little. It appears in retrospect that dimension 4 (time for planning) would be closely related to dimension 2 (home vs. away) in that time needed would logically depend on whether a breath freshener needed to be first acquired or not.

EXHIBIT 4

SITUATIONS BY FACTOR LOADINGS
AND
PRODUCTS BY UNSTANDARDIZED FACTOR SCORES

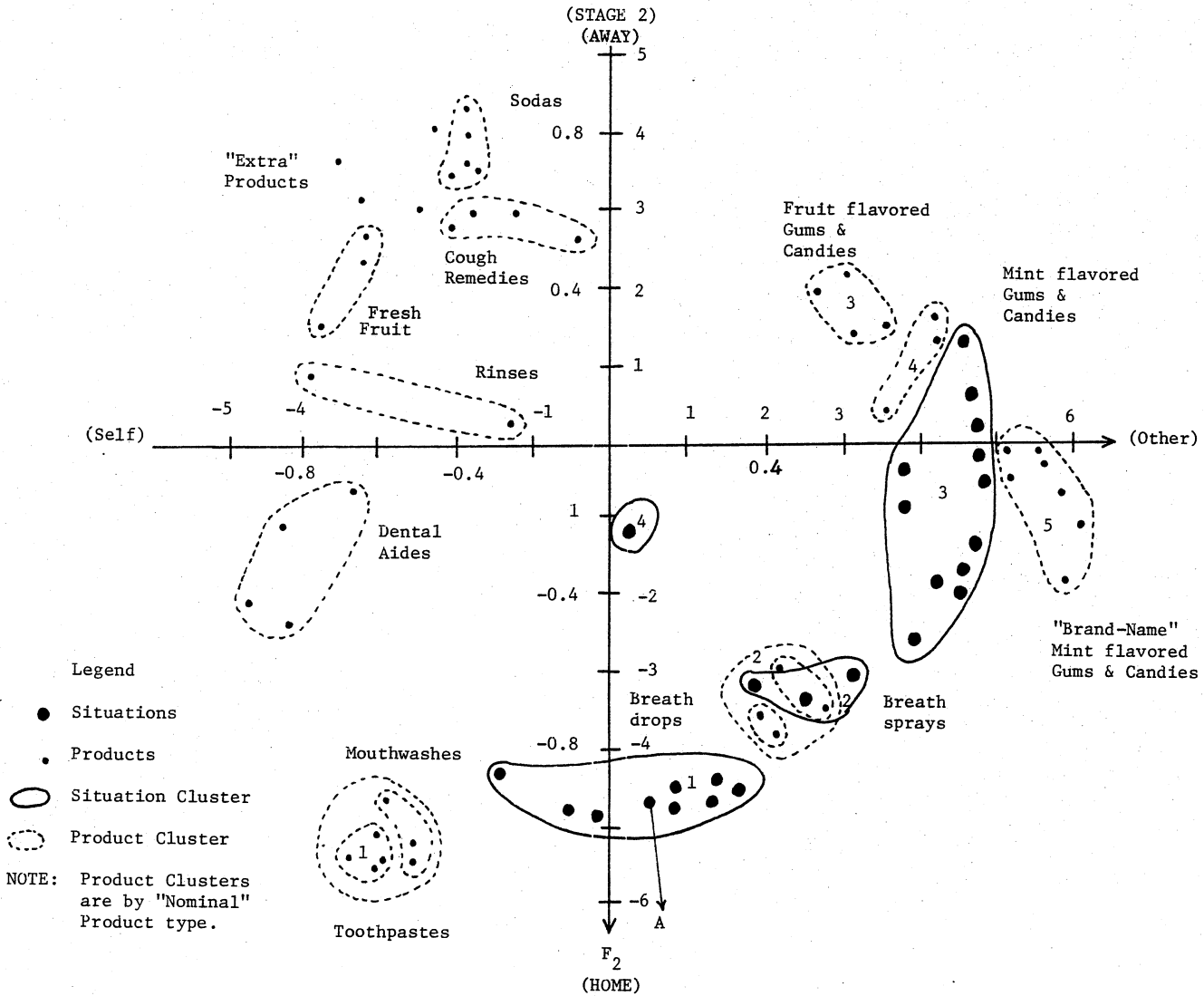


Exhibit 4 shows the 24 usage situations and 48 products of the second stage plotted in the space of the first two components. Points denoting the usage situations are merely plots of corresponding factor loadings (correlation coefficients with the principal components) which have directional properties only. The points denoting products are unstandardized factor scores, which have properties of both magnitude and direction.

Usage situation and product clusters were derived on the bases of patterns of factor loadings and factor scores, respectively, by visual clustering. In order to determine the relative appropriateness of product clusters for the situational clusters two methods could be used: (1) the projection of points representing products on the directional vector from the origin to the location of the centroid of each cluster of usage situations in reduced space would provide a measure of the relative appropriateness. The greater the projection along this directional vector, the more appropriate the product for the situational cluster. Thus in Exhibit 4, for situation "A", toothpastes and mouthwashes are considered more appropriate than breath

sprays and drops. This procedure may be repeated for each cluster of products and usage situations, (2) Alternatively, the original products-by-use matrix provides an explicit measure of appropriateness.

Situational clusters may be formed comprising situations that have similar product (cluster) appropriateness patterns and may be viewed as constituting a single market/submarket. This analysis revealed four submarkets (corresponding to the situational clusters shown in Exhibit 4).

For submarkets 2 and 4, product cluster 1 (mouthwashes and toothpastes) appears to compete with product group 5 (gums and candies). The extent of such competition, however, is an issue requiring further research.

Finally in order to assess the usefulness of the typology, average correlations were computed between pairs of situations coded in terms of the typology. These pairs had either zero to four typological dimensions in common (e.g., suppose situation A was coded 1, 0, 1, 1 (in terms of levels 1 or 2 of the four

typological dimensions) and situation B was coded 1, 0, 0, 1; then the two situations would have three dimensions in agreement). Correlations were across the original product appropriateness measures a_{ij} . The results are shown in Exhibit 5. As expected, as the number of typological elements in common increases, the extent to which pairs of situations have appropriate products in common also increases.

EXHIBIT 5

CONSISTENCY OF STAGE 2 USAGE SITUATIONS WITH TYPOLOGY

NUMBER OF DIMENSIONS in AGREEMENT	NO. OF PAIRS	AVERAGE CORRELATION	STANDARD DEVIATION
0	14	.25	.33
1	63	.36	.33
2	95	.47	.29
3	70	.72	.22
4	34	.84	.17

Discussion

Space did not permit the graphical representation results for Stage 1 (comparable to that reported for Stage 2 (see Exhibit 4)). If it had, the similarity of results between stages 1 and 2 would have been striking—this despite the several differences in stimuli (number of products, substantially different situational descriptions), the samples (eastern and western MBA candidates), and methods of data collection (card sorting versus partially overlapping questionnaires). All this is some testimony to the robustness of the procedures discussed.

As can be noted in Exhibit 4, brands within each generic product type (e.g., toothpastes, mouthwashes, breath sprays, cough drops) are generally close together in the space, indicating that subjects perceived fewer distinctions between brands than between product types and that such brands were generally perceived similarly in terms of their appropriateness to most usage situations (geometrically, the compactness of a cluster should connote a higher level of perceived substitutability with respect to all usage situations). It was particularly interesting to note, however, that flavor served to define the generic product rather than candies or gums. Cough drops were perceived differently from other candies (at least for breath freshening).

One particularly noteworthy consequence of our findings is the fact that just two principal components accounted for approximately 87 percent of the variation in association (93 percent for 3 components). This indicates that it may be possible to generate a parsimonious, yet meaningful situational typology for product types. Also, we observed that the first two factors in the two stages were quite similar. This may imply that "natural" situational descriptions serve about as well as morphologically-created ones for purposes of explaining systematic product-situation associations.

The relatively high measures of agreement among respondents would argue strongly against the presence of significant demand effects in the research. If respondents felt they should respond differently to each situation simply because of the task, it is unlikely that a_{ij} would be so high.

In stage 1, the appropriateness measure for each cell of the product-use matrix was based on 46 entries (respondents). On average, the number of entries of each cell was only 25 for stage 2. Yet, the results were strikingly similar. Again, this indicates the stability of the procedure for low ($n = 20$ to 40) sample sizes. A study by Rosenberg et al (1968) also reports stability of similarity matrices (based on concurrence/association measures) for small sample sizes ($n = 37$).

Conclusions

The results of the study are consistent with the following conclusions:

- (1) situational influences appear to affect the choice of products/brands in a systematic manner.
- (2) a parsimonious typology of objectively meaningful situations can be constructed based on behavioral criteria (the brands considered appropriate by consumers).
- (3) "ad hoc" (stage 1) descriptions of usage-situations serve about as well as descriptions based on a morphology (stage 2) in determining the composition of product-markets. In fact, the results of the two stages are strikingly similar.
- (4) product-market composition based on perceived consumer behavior across usage-situations would appear to stretch across generic product types (although the bases for defining a product-type may not always be apparent).
- (5) small sample sizes (30-40), using the procedure adopted, produced stable results.

Though we cannot claim to have developed a general situational taxonomy (although there is some similarity between our two dimensions and those of Belk (1975)), the procedures used here at least appear useful for a variety of scientific and managerial purposes. Further research is necessary with more significant sample sizes representing a broader cross-section of consumers and across a wider range of product-markets before generalizations can be claimed. But we have at least suggested that with product familiarity and usage situation controlled for, people's perceptions of the nature and composition of competitive markets are relatively stable. Individual differences in market behavior (e.g., in preferences and choice), which have become an important topic in consumer research, can more productively be explored within such a context.

References

- Allen, B.L. "Situational Factors in Conformity," in Leonard Berkowitz, Advances in Experimental Social Psychology, Vol. 2, New York: Academic Press, 1965.
- Barker, R.G. "Commentaries on Belk, 'Situational Variables and Consumer Behavior'". Journal of Consumer Research, Vol. 2, Dec. 1975, p. 165.
- Bearden, W.O. and A.G. Woodside "Interactions of Consumption Situations and Brand Attitudes," Journal of Applied Psychology, Vol. 61, No. 6, 1976, pp. 764-769.
- Belk, R.W. "Situational Variables and Consumer Behavior," Journal of Consumer Research, Vol. 2, Dec. 1975, pp. 157-164.
- Bourgeois, J.C., Haines, G.H. and M. Sommers "Defining an Industry," paper presented at the joint national conference of ORSA and TIMS, Las Vegas, Nevada: Nov. 17-19, 1975.
- Bowers, K.S. "Situationism in Psychology: An Analysis and Critique," Psychological Review. Vol. 80, No. 5, Sept. 1973, pp. 307-336,
- Day, G.S. and A.D. Shocker "Identifying Competitive Product Market Boundaries: Strategic and Analytical Issues." Working Paper #156, University of Pittsburgh, 1976.
- Endler, N.S. and D. Magnusson "Toward an Interactional Psychology of Personality," Psychological Bulletin, Vol. 83, No. 5, 1976, pp. 956-974.
- Frederiksen, N. "Toward a Taxonomy of Situations," American Psychologist, Feb. 1972, pp. 114-123.
- Hansen, F. "Psychological Theories of Consumer Choice," Journal of Consumer Research, Vol. 3, Dec. 1976, pp. 117-142.
- Jain, A.K. and M. Etgar "How to Improve Antitrust Policies with Marketing Research Tools," in Edward M. Mazze (Ed.), 1975 Combined Proceedings of the American Marketing Association, Series #37, 1975, pp. 72-75.
- Kruskal, J.B. "Relationship of Clustering to Scaling and the Interpretation of Neighborhoods," Bell-Penn Workshop on Multidimensional Scaling, June 7-10, 1972.
- Rosenberg, S., Nelson, C. and P.S. Vivekanathan "A Multidimensional Approach to the Structure of Personality Impressions," Journal of Personality and Social Psychology, Vol. 9, 1968, pp. 283-294.
- Russell, J.A. and A. Mehrabian "Environmental Variables in Consumer Research," Journal of Consumer Research, Vol. 3, June 1976, pp. 62-63.
- Steffire, V.J. New Products and New Enterprises: A Report of an Experiment in Applied Social Science, Irvine, Calif.: University of California, March, 1971.
- Wicker, A.W. "Processes Which Mediate Behavior - Environment Congruence," Behavioral Science, Vol. 17, May 1972, pp. 265-277.
- "Commentaries on Belk, 'Situational Variables and Consumer Behavior'", Journal of Consumer Research, Vol. 2, Dec. 1975, pp. 165-167.

ASSESSING THE EFFECTS OF VISIBLE
CONSUMPTION ON IMPRESSION FORMATION

Russell W. Belk, University of Illinois, Urbana-Champaign

Abstract

This paper reports a series of exploratory experiments measuring the impact of visible consumption of products and services on the impressions which others form about the consumers of these items. A method labeled "the detective study" is illustrated for measuring these impressions, and comparisons of effects are made between different types of consumption items and between male and female perceivers.

The belief that a person's possession and expenditures reveal something about the person may be one of the strongest cultural universals affecting consumer behavior. In virtually all cultures, visible products and services are the bases for inferences about the status, personality, and disposition of the owner or consumer of these goods. Relevant cues may be noted not only in the number and type of goods consumed, but also in such features as their style, color, uniqueness, condition and brand name. Where other information about a person is known, this information is integrated with visible consumption information in developing an overall impression of another. There are also many instances in which visible consumption cues dominate the overall impression, either because the cues are highly distinctive or because little additional personal information is available due to observing strangers or newly established or casual acquaintances. Nevertheless, while consumer behavior research has examined the messages about self or ideal self concept which consumers seemingly intend to display through their selections of products and services,¹ we have devoted little attention to the messages which consumption selections actually do convey to others. The broad intent of this research, therefore, was to initiate a program examining the influences which configurations of visible consumption items have on impression formation.

Prior Research

There has been little systematic research into the effects of visible consumption on impression formation, but there have been occasional studies examining isolated consumption factors affecting person perception. Perhaps the classic study was Haire's investigation of the image of Nescafé Instant Coffee, using subject descriptions of the presumed owner of a grocery shopping list including either instant or regular coffee (Haire, 1950). As indicated by subsequent replications, consumer images based on even these small differences in consumption patterns are significant and reliable as long as the product images remain constant (Westfall, Boyd, and Campbell, 1957), but may change as the images of the products change (Webster and Von Pechmann, 1970). A somewhat more visible item of consumption which has

received some attention in studies of person perception is clothing. Several researchers have found through unobtrusive experimental designs that those in higher status clothing (e.g., a suit and tie rather than casual clothes) seem to gain more compliance with their requests (e.g., signing a petition²) and also serve as models whose behavior is more imitated (e.g., crossing a street against a traffic light³). Holman (1976) has conducted an experimental study using photographs of a woman dressed in several different clothing ensembles, and found a number of differences in the inferred traits of the woman. While somewhat less conspicuous than clothing, cosmetics have also been found to affect person perception. In a study of the 1950's, McKeachie (1952) found that women wearing lipstick were judged by a sample of college men to be more frivolous, introspective, anxious, conscientious, and interested in the opposite sex than women not wearing lipstick. And Calder and Burnkrant (1977) recently found that a woman described as buying Revlon mascara was judged by college females to be more popular and competent than a woman described as buying Walgreens' mascara. Automobile ownership is another highly visible consumption characteristic which has been found to be related to person perception and related behaviors. Doob and Gross (1968) found that a more expensive and newer automobile which faked being stalled at a stoplight, received fewer horn-honking responses than a less expensive and older automobile. And Grubb and Hupp (1968) found that owners of Volkswagens and Pontiac GTOs were able to draw distinct and reasonably accurate profiles of the other group of owners' self concepts, knowing only the automobile which they owned.

In each of these studies, however, the investigated consumption items were examined in isolation from other consumption characteristics which might typically be apparent. It seems reasonable that people normally make judgments of others based on such broader configurations of visible consumption items. For instance, it is possible that a person who is wearing a suit and a tie with white stockings might be judged to be attempting to enact a role with which he is uncomfortable. In order to allow for such inferences, the present study sought a method which could systematically manipulate a composite of visible consumption items.

The "Detective Study"

The method chosen for the present research was to present each subject with a consumption profile representing a single cell from a factorial design. The experiment was presented in the guise of a "detective study" being done to help the New York City Police Department to more effectively locate the owners of unidentified property which was lost or stolen and later recovered by or turned in to the police. There are several

¹E.g., Grubb and Grathwohl (1967); Birdwell (1968); Grubb and Hupp (1968); Dolich (1969); Mason and Mayer (1970); Hamm and Cundiff (1969); Ross (1971); Grubb and Stern (1971); Greeno, Sommers, and Kernan (1973); and Landon (1974).

²Suedfeld, Bochner, and Metas (1971); Darley and Cooper (1972).

³Lefkowitz, Blake, and Mouton (1955).

advantages to this method. One desirable feature of the cover story is that it allows the presentation of a large number of consumption items without appearing forced. For instance, by including a wallet or purse in a particular group of items, their contents can also be manipulated. This allows the use of business cards, membership cards, credit cards, and specialty merchandise such as pocket calendars and matchbooks, all of which can manipulate the stimulus person's use of various moderately visible services. In addition, the use of this cover story removes any elements of social desirability which might otherwise inhibit the expression of consumption-related stereotypes. Subjects were told that the groups of items they would be shown in the descriptive form of "property tags" were ultimately returned to their owners, and that the characteristics of these owners were known. Thus, the stated purpose of the study was to compare perceptions of owners to the actual characteristics of the owners in order to see how much accuracy was possible from the items alone. This rationale also allowed the presentation of more than one group of unrelated items, amounting to participation in more than one experiment by the same subject.

Treatment Variables

All subjects were part of five different 2^5 factorial designs. There were ten subjects per cell for a total of 320 subjects participating in these experiments. Half of the subjects in each cell were males and all were undergraduate students at the University of Illinois. The five experiments differed in the visibility of consumption items included, in the sex of the stimulus person, and in whether or not the locational context where the items were reportedly found was manipulated. The two experiments which will be reported here both involved moderately visible consumption items for which locational context was not varied, but one involved a female stimulus person and the other involved a male stimulus person. The treatment levels for these two experiments are shown in Tables 1 and 2.

All treatment levels which a particular subject was to receive were simultaneously presented on 3 inch by 5 inch cards with the first two treatments presented on the same card and the order of the cards randomized. While it is not possible to claim that the two experiments were identical except for the sex of the stimulus person, both had a wallet or purse, a sports event, a restaurant, a personal care item, and either a travel mode or a travel abode. This rough equivalence allows comparisons between males judging either male or female stimulus persons, and females judging either male or female stimulus persons.

It was hypothesized that perceivers of the same sex as the stimulus person would be able to make more complete and consistent judgments about the person based on these visible consumption items (i.e., lower error variance) than could those of the opposite sex. The underlying assumption is that a person of the same sex will be able to bring more self-knowledge about use of the consumption items to bear on inferences about the stimulus person. For two of the consumption items (wallet/purse and haircare/lipstick) there may also be a greater familiarity with the nature of the consumption items themselves by a same sex observer. There has been a large amount of research on the effects of observer sex on person perception in studies not involving consumption stimuli,⁴ but the findings are inconsistent. The choice of the college subject pool, if it had an effect, probably biased results away from the hypotheses, since

⁴E.g., Taft (1955); Levy and Schlosberg (1960); Warr and Knapper (1968); and Tagiuri (1969).

at this prime age for dating, the students may be particularly alert to cues indicating the character of members of the opposite sex.

Descriptor Variables and Covariates

Based on pretests involving open-end responses describing the owners of many of the stimulus item configurations, the demographic and adjectival scales shown in Table 3 were selected as dependent variables ("descriptors"). These same variables, with the exception of the last two items, were used to obtain a self-description of the subject following participation in the five experiments. Because of the college subject pool, residence, family income, occupational class, and years of education were measured referring to subjects' parents.

The subject descriptions will serve as covariates in future analyses, but at present it is of interest to compare the intercorrelations of these covariates, the intercorrelations of the descriptor variables, and the cross correlations between covariates and descriptor variables, in order to consider whether several types of response styles may have affected judgments. This information is summarized in Figure 1 for the moderately visible male consumption items experiment using data from the 160 male subjects.⁵ An examination of the three matrices in Figure 1 allows several types of potential insights into subject responses. If the configurations of the covariate intercorrelations and the descriptor intercorrelations were equivalent, this could be taken as evidence of inherent clusters of traits, and factor or cluster analysis might be used to reduce the dependent measures prior to further analysis. Although there is some similarity between these two matrices, the differences in the number of significant correlations forces rejection of this conclusion and course of action. If on the other hand, the covariate intercorrelations were stronger than the descriptor intercorrelations, then one possibility would be that "halo effects" exist in self ratings and stimulus persons are viewed more discriminately. Another possible interpretation of such a pattern would be that while certain traits naturally occur together in the subjects (e.g., educated/successful), the consumption items presented are unnatural combinations which cause normal trait associations to fall apart. Since this pattern did not emerge, neither of these interpretations need concern us. A third possible pattern, and the one which is found in the data in Figure 1, is that the covariate intercorrelations will be weaker than the descriptor intercorrelations. The explanation in this case is clearer. The tendency to judge others using bundles of related attribute judgments has been referred to as "cojudgment" in person perception studies not involving consumption items (Warr and Sims, 1965). This phenomenon appears to be due to a reliance on implicit theories of personality types when judging others. That is, when judging others we tend to rely upon our knowledge or impressions of people in general to go beyond surface traits in deciding what characteristics people are likely to possess. This tendency leads to the use of groups of normally related personality characteristics. An overlapping but less plausible explanation for this pattern is that consumption characteristics reveal patterns of personality characteristic relationships which would not normally be found together. For instance, possession of a Rolls Royce automobile

⁵The absolute magnitudes and patterns of correlations differed somewhat by subject sex and stimulus person sex, but the relative strengths of the covariate intercorrelations versus the "projection diagonal" correlations versus the descriptor intercorrelations remained the same ordering over the four combinations of subject and stimulus person genders.

TABLE 1

Treatment Levels* for Moderately Visible Male Consumption Items Experiment

- Treatment 1: "A slightly worn slimline man's wallet of brown...
(a) Moroccan handstitched leather"
(b) simulated cowhide"
- Treatment 2: "...found in the (a) United Airlines portion of New York's JFK Airport"
(b) Greyhound Bus Terminal in Manhattan"

"...containing..."
- Treatment 3: "A pocket calendar from the (a) Valley Beef and Ale House, N.Y.C."
(b) Valli Chinese Restaurant, N.Y.C."
- Treatment 4: "A business card for (a) Frank's Barber Shop, N.Y.C."
(b) Unisex Hairstyling Boutique, N.Y.C."
- Treatment 5: "Two tickets for a (a) New York Jets home football game"
(b) New York Mets home baseball game"

* All treatments at two levels, (a) and (b).

TABLE 2

Treatment Levels* for Moderately Visible Female Consumption Items Experiment

- Treatment 1: "A relatively new black leather (a) shoulder bag"
(b) hand bag"
- Treatment 2: "...found in New York's Madison Square Garden following a championship (a) hockey game"
(b) tennis
tournament"

"...containing..."
- Treatment 3: "A matchbook from King's (a) Cafeteria, N.Y.C."
(b) Restaurant, N.Y.C."
- Treatment 4: "A green plastic ballpoint pen inscribed with the name (a) 'Holiday Inn Motel'"
(b) 'Waldorf Astoria Hotel'"
- Treatment 5: "A tube of (a) bright red 'Revlon' lipstick"
(b) pale pink 'Revlon' lipstick"

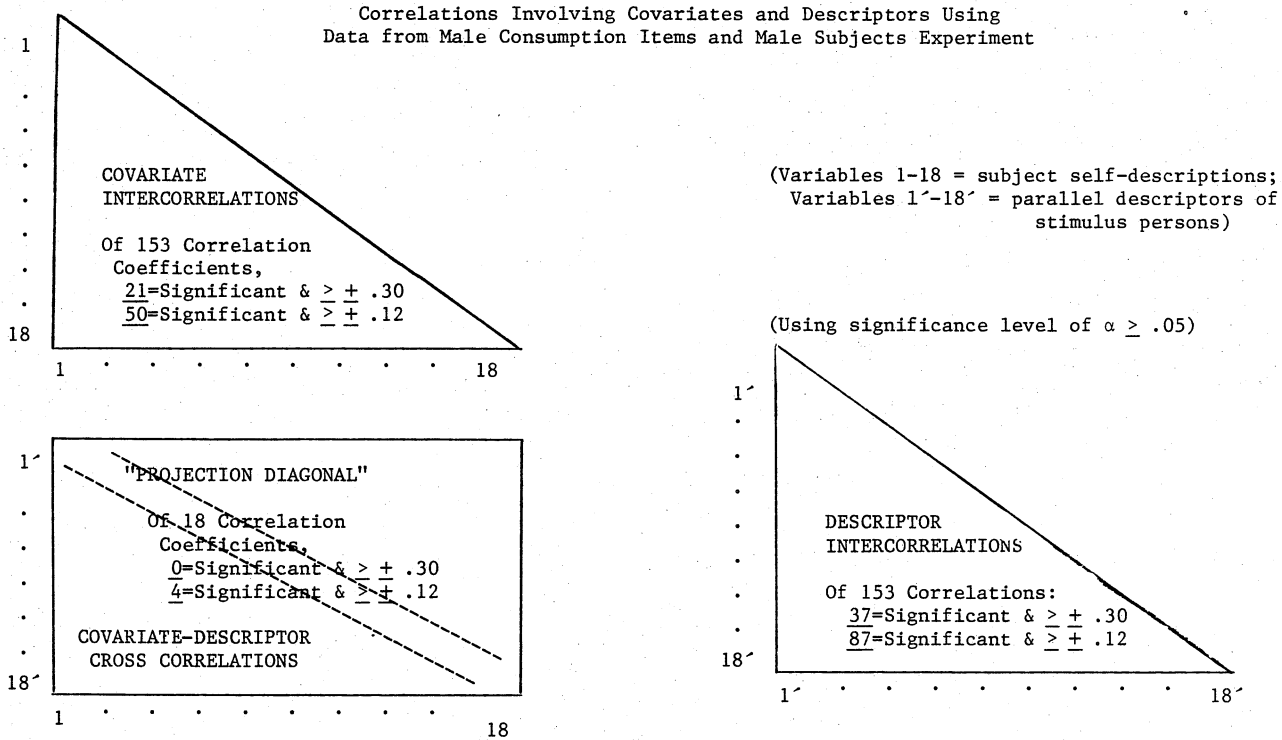
* All treatments at two levels, (a) and (b).

TABLE 3

Dependent Measures

- | | |
|-----------------------------------------------|--------------------------------------------------------------------|
| 1. Family Life Cycle Stage | 12. Friendly...Unfriendly |
| 2. Age | 13. Likeable...Unlikeable |
| 3. Residence (Urban, Suburban, Rural) | 14. Attractive...Unattractive |
| 4. Family Income | 15. Emotional...Unemotional |
| 5. Occupational Class (9 Prestige Categories) | 16. Successful...Unsuccessful |
| 6. Years of Education | 17. Aggressive...Passive |
| 7. Generous...Stingy (7-Point Bipolar Scale) | 18. Interesting...Dull |
| 8. Serious...Fun-Loving | 19. Sex (No Variance for Experiments to Be Reported) |
| 9. Responsible...Irresponsible | 20. Amount of Money (if any) Missing |
| 10. Happy...Unhappy | 21. Verbal Description of Any Other Impressions
of Items' Owner |
| 11. Heavy...Thin | |

FIGURE 1



might lead to judgments that the owner is of high income and also ostentacious, even though the two characteristics are not normally correlated. However, if we assume that an observer tries to not only interpret the characteristics of a stimulus person, but also to form a familiar interpretation which will allow further predictions of that person's demeanor and behavior, then the cojudgment interpretation of descriptor intercorrelations is the most plausible.

One further possible basis for judgments of stimulus persons is examined in the "projection diagonal" of Figure 1. These are the correlations between the corresponding covariate and descriptor items. High positive correlations here would indicate that the subjects tend to project their own perceived traits onto the stimulus person. As the weak correlations indicate, this projection does not appear to have occurred. Instead it seems that while there may be some basis for expecting inherent trait clusters, cojudgment occurs in judging others to a greater extent than it does in self-descriptions. From the point of view of data analysis, this indicates that data reduction in the set of dependent measures should not be expected to yield a set of underlying variates which represent general person traits common to stimulus persons as well as subjects of both sexes. Avoiding data reduction also eliminates the potential problem of averaging the differences in relationships which may occur across cells of the experimental design.

Results

Because of the strong likelihood that multivariate analyses of variance would form different linear composites of the dependent variables across the two experiments and two subject groups of interest, MANOVA procedures were rejected for the present comparisons. While it would be possible to include subject sex as a sixth

independent variable in univariate analysis of variance, the complicated potential interactions between sex and the treatment variables would prove difficult to interpret. Therefore, the method of analysis chosen was to run separate ANOVAs for each dependent variable in each of the four combinations: (1) male subjects, male stimulus person experiment; (2) male subjects, female stimulus person experiment; (3) female subjects, male stimulus person experiment; and (4) female subjects, female stimulus person experiment. A sample ANOVA for the dependent variable "unattractiveness" with the data from combination 1, is presented in Table 4.

It may be seen from Table 4 that three main effects and three interactions are significant. The way in which these variables affect attractiveness judgments will be described shortly. For the moment, it is more important to note that these effects were able to account for 62 percent of the variance in attractiveness ratings, based on variance components estimates.⁶ This is the measure which is to be used for testing the hypothesis that subjects will be able to judge those of the same sex more completely and consistently based on consumption items. Table 5 shows a summary of proportions of variance accounted for in each dependent variable for the four subject/stimulus person combinations.

Looking only at the average proportions of variance accounted for over all dependent measures, it appears that females are clearly better judges of both male and female consumers than are males. Female judgments were more consistent across judges, leading to higher levels of explained variance. Male and female stimulus persons, on the other hand, appear to be equally easy to judge. This is true whether the perceiver is male

⁶See Dwyer (1974).

TABLE 4

ANOVA for Unattractiveness Variable for Male
Subjects in Male Stimulus Person Experiment

Source	Sum of Squares	D.F.	Mean Square	F	Prob.
A Airport/Bus Terminal	.15	1	.15	.16	--
B Leather/Simulated Wallet	22.26	1	22.26	24.32	.001
C Beef/Chinese Restaurant	3.91	1	3.91	4.09	.043
D Barber/Hairstylist	33.31	1	33.31	34.83	.001
E N.Y. Jets/N.Y. Mets	1.06	1	1.06	1.11	--
B x D	3.91	1	3.91	4.09	.043
Other 2-way interactions	7.36	9	.82	.96	--
A x B x D	6.81	1	6.81	7.12	.008
A x B x E	5.26	1	5.26	5.50	.020
Other 3-way interactions	6.30	8	.79	.82	--
4-way interactions	2.98	5	.60	.62	--
5-way interaction	1.06	1	1.06	1.11	--
Error	122.40	128	.96		
Total	217.74	159			

Proportion of Variance Accounted for (Fixed Effects Model): .62

TABLE 5

Proportions of Variance Accounted for in Each Dependent
Variable for Four Combinations of Subject and Stimulus Person Gender

Dependent Variable	Experiment (Subject Gender/Stimulus Person Gender)			
	Male/Male	Female/Male	Male/Female	Female/Female
Life Cycle Stage	.26	.24	.35	.42
Age	.62	.35	.41	.48
Residence	.65	.36	.36	.62
Family Income	.30	.49	.12	.45
Occupation	.59	.40	.19	.47
Generous	.17	.29	.54	.45
Serious	.44	.46	.41	.46
Responsible	.47	.39	.51	.49
Happy	.24	.53	.39	.33
Heavy	.40	.50	.35	.43
Friendly	.28	.49	.33	.57
Likeable	.11	.66	.33	.30
Attractive	.62	.61	.28	.56
Emotional	.06	.09	.46	.42
Successful	.31	.61	.37	.30
Aggressive	.52	.64	.50	.38
Interesting	.42	.67	.41	.48
Money Missing	.38	.62	.53	.46
Average Proportion	.38	.47	.38	.45

or female. Examining the proportions of variance accounted for in each separate independent measure, it appears that females are consistently the best judges of income, seriousness, weight, friendliness, and interestingness, while males more readily assess responsibility. Judges of the same sex as the stimulus person seem to do somewhat better at estimating age, life cycle stage, residence, occupation, and attractiveness, while judges of the opposite sex as the stimulus person were able to more consistently estimate generosity, happiness, likeableness, emotionality, successfulness, aggressiveness, and amount of money missing, using consumption characteristics. Generally, however, it must

be concluded that these patterns of differences fail to conform to an easily identifiable sensitivity stereotype such as one sex being more able to agree on affective judgments while the other sex focuses on more objective or cognitive characteristics.

Although the hypothesized ability of perceivers to judge those of the same sex more completely based on their consumption patterns was not supported in total variance explained, it is still possible that such effects are reflected in male and female differences in their interpretations of a particular consumption characteristic. In order to assess this possibility,

Tables 6 and 7 provide a summary of the direction of main effects found to be significant ($\alpha \leq .05$) in the two experiments, broken down by sex of the perceiver. Although space precludes presenting interactions, a substantial portion of the variance in these two experiments was accounted for by these main effects. In Table 6 and to an even greater degree in Table 7 it may be seen that the implications of these consumption item clues differs between male and female subjects. For the wallet and hand or shoulder bag items it does

appear that the subjects presumed to be personally most familiar with these items were able to provide more consistent descriptions of their owners. While this lends support to the hypothesized bases for sex effects in impressions formed from visible consumption, the same pattern does not emerge for the lipstick and haircare items.

TABLE 6

Summary of Main Effects for Moderately Visible Male Consumption Items Experiment

Treatment 1: Those whose wallets were found in an air terminal rather than a bus terminal were perceived...

<u>By Males Only As:</u>	<u>By Both Males and Females As:</u>	<u>By Females Only As:</u>
	Higher Income	Older
	Higher Occupational Class	More Generous
	More Highly Educated	More Responsible
	More Likeable	Friendlier
	More Successful	More Attractive
	More Interesting	More Aggressive
		Missing More Money

Treatment 2: Those whose wallets were handstitched Moroccan leather rather than simulated cowhide were perceived...

<u>By Males Only As:</u>	<u>By Both Males and Females As:</u>	<u>By Females Only As:</u>
More Fun-Loving	Higher Income	
Happier	Higher Occupational Class	
Thinner	More Highly Educated	
Friendlier	More Successful	
More Likeable	More Aggressive	
More Attractive		
More Interesting		
Missing More Money		

Treatment 3: Those whose wallets contained a calendar from a beef and ale house rather than from a Chinese restaurant were perceived...

<u>By Males Only As:</u>	<u>By Both Males and Females As:</u>	<u>By Females Only As:</u>
Less Attractive	More Likely to Be Suburban	More Likeable
More Aggressive		

Treatment 4: Those whose wallets contained a business card from a barber rather than from a hairstylist were perceived...

<u>By Males Only As:</u>	<u>By Both Males and Females As:</u>	<u>By Females Only As:</u>
Lower Occupation	Older	More Suburban
Less Highly Educated	More Likely to Be Married	More Responsible
	Heavier	Less Interesting
	Less Attractive	Missing Less Money

Treatment 5: Those whose wallets contained tickets to a football game rather than tickets to a baseball game were perceived...

<u>By Males Only As:</u>	<u>By Both Males and Females As:</u>	<u>By Females Only As:</u>
Higher Occupation		More Fun-Loving
More Successful		Friendlier
		More Interesting

TABLE 7

Summary of Main Effects for Moderately
Visible Female Consumption Items Experiment

Treatment 1: Those whose bags were found at a hockey game rather than at a tennis tournament were perceived...

<u>By Males Only As:</u>	<u>By Both Females and Males As:</u>	<u>By Females Only As:</u>
Heavier		
Less Successful		
More Suburban		

Treatment 2: Those with shoulder bags rather than hand bags were perceived...

<u>By Males Only As:</u>	<u>By Both Females and Males As:</u>	<u>By Females Only As:</u>
Missing Less Money	Older	Lower Income
		Lower Occupation
		Less Education
		Heavier
		Less Attractive
		Less Successful
		Less Interesting

Treatment 3: Those with a matchbook from a cafeteria rather than from a restaurant were perceived...

<u>By Males Only As:</u>	<u>By Both Females and Males As:</u>	<u>By Females Only As:</u>
Younger		More Emotional
Less Educated		
Less Successful		
Less Responsible		
Less Aggressive		
Less Interesting		
Missing Less Money		

Treatment 4: Those with pens from a Holiday Inn rather than from the Waldorf Astoria Hotel were perceived...

<u>By Males Only As:</u>	<u>By Both Females and Males As:</u>	<u>By Females Only As:</u>
Less Educated	Younger	More Likely Married
Medium Occupation	Lower Income	Less Responsible
More Fun-Loving		Less Emotional
Missing Less Money		Less Successful

Treatment 5: Those with a tube of red lipstick rather than pink lipstick were perceived...

<u>By Males Only As:</u>	<u>By Both Females and Males As:</u>	<u>By Females Only As:</u>
Older		
Less Interesting		

Discussion

For the two experiments for which results were presented, given the confidence level used in testing, nineteen main effects could be expected to occur by chance over the two subject groups. In fact, more than five times this many main effects were found to be significant. Even moderately visible and sometimes subtle consumption differences were found to produce consistent differences in the impressions formed of the consumers of these products and services. Of course, the particular effects obtained in these experiments can be expected to change with different groups of perceivers and over time. It is also likely that effects may differ somewhat with different presentation contexts. For instance, if consumption items are presented via photographs or videotapes, the cue that these items were lost or stolen disappears, the physical features of the stimulus person are introduced, and selective perception is likely to be more operative. Nevertheless, the detective study methodology presents an efficient way to examine the combined effects of a number of visible consumption items on impression

formation.

An examination of the effects summarized in Tables 6 and 7 reveals one disturbing but expected result. For example, consider the effect of the air/bus treatment in Table 6 on females. Not only was the air traveler judged by relatively objective wealth-related criteria as being higher income, higher occupational status, more highly educated, and missing more money than the bus traveler, but he was judged by criteria objectively unrelated to wealth to be more likeable, successful, interesting, generous, responsible, attractive, and aggressive! In part, this halo effect may be due to the business student subjects' identification with the socially distant reference group of executive/air travelers,⁷ but it also seems apparent that the more objective judgment of wealth creates a bias toward an array

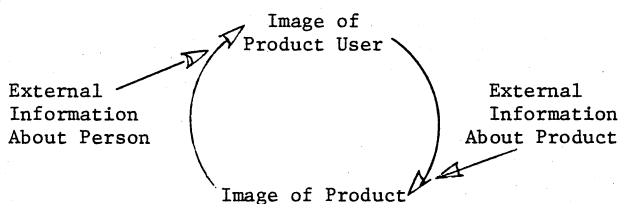
⁷ See Cocanougher and Bruce (1971).

of other positive characteristics which bear no logical relationships to income. Whether estimates of income are in fact focal inferences which bring with them a number of collateral judgments in forming impressions based on visible consumption is a question which must await more detailed examinations of information processing and integration of such cues. Two somewhat parallel processes which may shed some light on this question are investigations of the integration of brand attribute information into an overall brand attitude, and the integration of information about the products and brands carried by a retailer into the overall impressions of the retailer.

We need to know much more about the role of visible consumption in impression formation. As Figure 2 illustrates, impression formation is an interactive process involving images of products and services as well as images of people consuming these items. Inferences about a product user are determined in part by the products he or she is seen to use or consume, provided that some prior image of these products exists for the perceiver. Similarly, inferences about the image of a product are partially determined by those seen to use or consume the product, provided that some prior image of these users exists for the perceiver. In addition to previously received external information about the person or product, inferences may be aided by information about the consumption situation in which a product is being used (e.g., an outfit being worn in a church rather than on a tennis court), and information about the role being enacted by the consumer (e.g., wearing an outfit as a salesperson in a particular store rather than as a customer in this store). The effects of product image on inferences about product user image are likely to be strongest where little prior information about the person exists, but Bem (1967) argues that some such inferences may take place even when the perceiver is the stimulus person. Whether or not such self-perception is normally a part of impression formation via visible consumption, however, it is clear that in order to know more about the role and function of product images, we need to know more about how the use of visible products and services influences our impressions of people.

FIGURE 2

Sources of Information Affecting Interpretations of People and Products



References

- Darrel J. Bem, "Self Perception: An Alternative Interpretation of Cognitive Dissonance Phenomena," Psychological Review, 74 (1967), 183-200.
- Al E. Birdwell, "A Study of the Influence of Image Congruence on Consumer Choice," Journal of Business, 41 (1968), 76-88.
- Bobby J. Calder and Robert E. Burnkrant, "Interpersonal Influence on Consumer Behavior: An Attribution Theory Approach," Journal of Consumer Research, 4 (1974), 29-38.
- A. Benton Cocanougher and Grady D. Bruce, "Socially Distant Reference Groups and Consumer Aspirations," Journal of Marketing Research, 8 (1971), 379-381.
- John M. Darley and Joel Cooper, "The 'Clean for Gene' Phenomenon: The Effect of Students' Appearance on Political Campaigning," Journal of Applied Social Psychology, 2 (1972), 24-33.
- Ira J. Dolich, "Congruence Relationships Between Self Images and Product Brands," Journal of Marketing Research, 6 (1969), 80-84.
- Anthony N. Doob and Allen E. Gross, "Status of Frustrator as an Inhibitor of Horn-Honking Responses," Journal of Social Psychology, 76 (1968), 213-218.
- James Dwyer, "Analysis of Variance and the Magnitude of Effects: A General Approach," Psychological Bulletin, 81 (1974), 231-237.
- Daniel Greeno, Montrose S. Sommers, and Jerome B. Kernan, "Personality and Implicit Behavior Patterns," Journal of Marketing Research, 10 (1973), 63-69.
- Edward L. Grub and Harrison L. Grathwohl, "Consumer Self-Concept, Symbolism, and Market Behavior: A Theoretical Approach," Journal of Marketing, 5 (1967), 22-27.
- Edward L. Grubb and Gregg Hupp, "Perception of Self, Generalized Stereotypes, and Brand Selection," Journal of Marketing Research, 5 (1968), 58-63.
- Edward L. Grubb and Bruce L. Stern, "Self-Concept and Significant Others," Journal of Marketing Research, 8 (1971), 382-85.
- Mason Haire, "Projective Techniques in Marketing Research," Journal of Marketing, 14 (1950), 649-656.
- Bruce C. Hamm and Edward W. Cundiff, "Self-Actualization and Product Perception," Journal of Marketing Research, 6 (1969), 470-472.
- Rebecca C. Holman, "Communicational Properties of Women's Clothing: Isolation of Discriminable Clothing Ensembles and Identification of Attributions Made to One Person Wearing Each Ensemble," unpublished dissertation, University of Texas, 1976.
- E. Laird Landon, Jr. "Self-Concept, Ideal Self-Concept, and Consumer Purchase Intentions," Journal of Consumer Research, 1 (1974), 44-51.
- Monroe Lefkowitz, Robert R. Blake, and Jane S. Mouton, "Status Factors in Pedestrian Violation of Traffic Signals," Journal of Abnormal and Social Psychology, 51 (1955), 704-706.

Nissum Levy and Harold Schlosberg, "Woodworth Scale Values of the Lightfoot Pictures of Facial Expression," Journal of Experimental Psychology, 60 (1960), 121-125.

Joseph B. Mason and Morris L. Mayer, "The Problems of the Self-Concept in Store Image Studies," Journal of Marketing, 34 (1970), 67-69.

William J. McKeachie, "Lipstick as a Determinator of Personality," Journal of Social Psychology, 36 (1952), 241-244.

Ivan Ross, "Self-Concept and Brand Preference," Journal of Business, 44 (1971), 38-50.

Peter Suedfeld, Stephen Bochner, and Carol Metas, "Petitioner's Attire and Petition Signing by Peace Demonstrators" A Field Experiment," Journal of Applied Social Psychology, 1 (1971), 278-283.

Ronald R. Taft, "The Ability to Judge People," Psychological Bulletin, 52 (1955), 1-23.

Renato Tagiuri, "Person Perception," in Gardner Lindzey and Elliot Aronson (eds.), Handbook of Social Psychology, 2nd Edition, Volume 3 (Reading: Addison-Wesley, 1969), 395-449.

Peter B. Warr and Christopher Knapper, The Perception of People and Events (London: Wiley, 1968).

Peter B. Warr and Alan Sims, "A Study of the Cojudgment Process," Journal of Personality, 33 (1965), 598-604.

Frederick E. Webster, Jr. and Frederick von Pechmann, "A Replication of the 'Shopping List' Study," Journal of Marketing, 34 (1970), 61-63.

Robert L. Westfall, Harper W. Boyd, and Donald T. Campbell, "The Use of Structured Techniques in Motivation Research," Journal of Marketing, 22 (1957), 134-139.

EVOKED SET SIZE AS A FUNCTION OF NUMBER OF CHOICE CRITERIA AND INFORMATION VARIABILITY

Joseph A. Belonax, Jr., Bowling Green State University
Robert A. Mittelstaedt, University of Nebraska - Lincoln

Abstract

If consumers construct an evoked set by a sorting process, applying more choice criteria to information which contains more variability should increase the likelihood of achieving a smaller evoked set. A 2 X 3 factorial experiment to test these hypotheses is reported. Results support both hypotheses.

Problem

Conceptualizing the consumer as an information processor and decision maker, Chaffee and McLeod (1973) present a generalized model which defines the area of consumer information processing. The consumer is seen as making a buying decision according to some rule which combines information from an attribute-by-brand matrix. Within the context of this broad model, studies of information processing may be classified as structural or functional in their orientations.

Structural issues which have been investigated include the nature and number of product attributes used in making product evaluations and/or choices (a number of such studies are cited and discussed in Myers and Alpert, 1977). Other studies have examined the nature of the "evoked set," i.e., the number of brands actually considered by consumers when making a choice (Campbell, 1969; Gronhaug, 1973/74; Jarvis and Wilcox, 1973; May and Homans, 1977; Narayana and Markin, 1975; Ostlund, 1973).

Research on functional aspects of information processing has taken many forms (summarized in Bettman, 1977) but often involves fitting one or more decisions rule models to brand evaluations made by persons who have been exposed to an attribute-by-brand matrix of information. As Bettman (1977) has pointed out, some of these models (affect referral, linear compensatory, disjunctive and conjunctive) assume that persons process information by brand while others (lexicographic, elimination by aspects, semi-order lexicographic and additive difference) assume processing by attribute. The somewhat mixed results of studies attempting to fit various models may reflect situational differences in those studies. Further, as Bettman notes, modes of information presentation may bias results in favor of one model or another.

However, coming from a different perspective, it may be suggested that the attempts to fit a single model to a total decision process, especially when the buying situation has one or more relatively unfamiliar aspects, may be overlooking an important intermediate stage. The concept of evoked set suggests that choices are made only after the consumer has constructed a set of "acceptable" brands (Howard and Sheth, 1969). The implied two-process model, with information first used to form an evoked set, is consistent with the results reported by Pras and Summers (1975). They found the conjunctive model to be the best predictor of preference rank order when all alternatives were considered, but the linear compensatory model to predict best when only acceptable alternatives were considered.

This suggests that the construction of an evoked set may be an important phenomenon and this study is addressed to that issue. While no decision rule models will be fit to preference orderings, the number of alternatives

considered acceptable (evoked set size) will be examined with respect to two variables -- the number of choice criteria employed and the nature of the information presented -- according to hypotheses which follow from a "sorting view" of evoked set construction.

Hypotheses

Although all consumers may be exposed to the same set of alternative brands in an experimental setting or in the market place, the concept of evoked set implies that not everyone makes a choice from the same set of brands. Indeed, there appear to be substantial inter-individual differences in evoked set size for particular products (Campbell, 1969; Gronhaug, 1973/74; Narayana and Markin, 1975). Confronted with an array of information about a fairly large number of brands, it seems reasonable to suppose that a person would try to simplify the situation. While this could be done by reducing the number of brands or reducing the number of choice criteria, it seems more likely that a person would try to reduce the number of brands actively considered for at least two reasons. First, it would reduce the number of comparisons which would ultimately have to be made. Second, and assuming we are talking about attributes of importance, to simplify by eliminating choice criteria would force the consumer to, in effect, forego some of the consumption goals which he or she is trying to accomplish through the purchase of the product.

The consumer's task, then, is to eliminate brands from active consideration based on information about attributes of importance. Since the purpose of evoked set construction is to simplify the ultimate choice process, it seems likely that a person would use an "elimination rule" and proceed by "knocking out" brands which failed to meet some acceptable minimum level on one or more evaluative criteria. It would follow that, as the consumer applied a larger number of evaluative criteria to the awareness set, the likelihood of finding a reason for eliminating brands would increase. Further, the nature of available information should play a role. If the information about attributes consists of some sort of ratings of alternative brands, as the variability of those ratings increased and/or the general level of their desirability decreased, the likelihood of finding a reason to eliminate brands should also increase. Therefore, specifically hypothesized in this study are the following:

Hypothesis 1: The size of the evoked set is inversely related to the number of choice criteria used in the evaluation task.

Hypothesis 2: The size of the evoked set is inversely related to the variability and directly related to the mean level of the attribute ratings presented in the evaluation task.

Method

To test these hypotheses, a 2 x 3 factorial experiment was designed with "number of evaluative criteria used" and "variability and mean levels of presented attribute information" as independent variables and the reported size of the evoked set as the dependent variable. Before discussing the operationalization of the variables,

four preliminary design considerations will be discussed.

First, the product for which brand information was to be supplied had to be sufficiently complex to allow the use of several evaluative criteria. Further, as will be argued later, it was desirable to select a product about which participants would be relatively uninformed and, yet, one in which they would be somewhat interested. A product which appeared to meet these criteria, micro-wave ovens, was selected.

Second, since the number of attributes was an independent variable, it was necessary to attempt to insure that brand rating information was presented for attributes of importance. Forty-eight knowledgeable respondents (senior home-economics majors enrolled in a course which had previously discussed micro-wave ovens) rated the importance of and perceived interbrand differences for 16 attributes. Using the "direct dual questioning technique" described by Alpert (1971), 6 attributes with average importance and difference ratings significantly ($p < .01$) greater than the grand mean were selected as "salient." The six included browning uniformity, cooking precision, cooking uniformity, cooking versatility, oven capacity and warranty protection. When presented to participants in the experiment, the nature of each attribute was described in a short paragraph. Participants began the evaluation task by reading those descriptions which were then briefly discussed by the experimenter.

Third, since the size of the evoked set appears to be related to the size of the awareness set (Jarvis and Wilcox, 1973), it was necessary to hold constant the number of brands about which information was provided. To manipulate the variability of the attribute ratings "standard squares" were used (for reasons described below.) Since there were six salient attributes, information was presented on six "brands" to each participant.

Fourth, because other studies have suggested a relationship between product and/or brand familiarity and the size of the evoked set (Gronhaug, 1973/74; Jarvis and Wilcox, 1973; Ostlund, 1973), it was necessary to try to control for product familiarity. To this end, participants for the study were 300 students in an introductory nutrition course which had not discussed micro-wave ovens in class. Random assignment of participants to the several treatment levels should have reduced the effects of prior product experience. Finally, the "brands" were presented as "real, but disguised" and designated by the letters "A" through "F."

One independent variable, the variability and mean level of attribute ratings, was manipulated by systematically changing the ratings of the 6 previously determined "salient" attributes for each of the six "brands." Within each treatment level, the mean and variability were held constant by the use of "standard squares" (Kirk, 1968). However, between treatment levels, both the mean level and variability were systematically changed from the "low variability" condition ($X = 8.0$; $s^2 = 5.9$), through the "intermediate variability" condition ($X = 7.0$; $s^2 = 8.4$) to the "high variability" condition ($X = 6.0$; $s^2 = 14.0$). It will be noted that, to provide information which could be used for "sorting," the mean levels were decreased as the variability was increased.

While it is possible to hold the means constant and manipulate the variability, such a design would necessitate choosing a rather "middling" value for the mean and would result, in the lower variability conditions, in fewer (if any) "high" ratings on individual attributes. Since some persons are likely to focus on those attributes with high ratings, manipulating variability

without the simultaneous adjustment of mean values runs a serious risk of having a "self cancelling" effect.

The other independent variable was a stratification according to the number of choice criteria used. After examining the brand ratings and making an overall evaluation of each brand, participants were asked to indicate: (a) the importance of each attribute to the evaluation decision, (b) the minimum rating on each attribute which an oven would have to receive to be evaluated as "acceptable" and (c) the number (and designating letters) of brands considered to be acceptable for purchase. (The answer to this last question is the dependent variable, the evoked set size.) Answers from questions (a) and (b) were combined and a mean calculated for each attribute. Any participant's combined answer on any attribute which exceeded the mean for that attribute was designated an "employed choice criterion" and the number of such was determined for each participant. Previously it had been determined to consider those using 1, 2 or 3 criteria as the "few choice criteria" group and those using 4, 5 or 6 as the "many" choice criteria group. Splitting in this way resulted in reasonably equal numbers in each condition, with 164 in the "few" group and 136 in the "many" category.

Since the other independent variable was randomized across the entire group of participants, the two variables are "crossed" and the design factorial. Data were analyzed using the ANOVA subroutine of MULTIVARIATE (Finn, 1972). Although there were some differences in cell size (Table 1), reordering the variables for unequal cells did not affect significance.

Results

Table 1 shows the mean evoked set size for the various conditions of the study. An examination of the means

TABLE 1
Mean Evoked Set Size For All Conditions

Choice Criteria	Attribute Rating Variability			
	Low	Intermediate	High	Total
Many	2.55 (58)*	2.36 (39)	1.79 (39)	2.28 (136)
Few	3.09 (44)	2.76 (59)	2.18 (61)	2.63 (164)
All	2.78 (102)	2.60 (98)	2.03 (100)	2.45 (300)

* ("n" in parentheses)

shows that, as predicted in Hypothesis 1, those employing "few" choice criteria constructed larger evoked sets than those employing "many" choice criteria. Hypothesis 2 predicted that the evoked set would be smaller as the variability of the presented attribute ratings increased while their mean levels decreased. The means are in the predicted order for both the "many" and the "few" choice criteria groups and, therefore, in total. Table 2 presents the results of the two-way univariate analysis of variance. Both main effects are significant while the interaction is not. The results confirm both hypotheses.

TABLE 2
Univariate Analysis of Variance

Source of Variation	d. f.	Mean Square	F
Choice Criteria	1	9.356	10.002 (p<.01)
Rating Variability	2	18.038	19.284 (p<.01)
Interaction	2	.173	.185 (NS)
Error	294	.936	

Discussion

The results confirm the view that, even though relatively few brands were presented for evaluation, participants did not consider all of them to be acceptable. The "grand mean" is an evoked set size of only 2.45 brands. That the size of the evoked set seems to depend on the number of choice criteria used and the nature of the available information is consistent with the expectation that participants would engage in a simplification strategy to construct an evoked set. The evoked set is largest when relatively few choice criteria are applied to information which indicates that the brands are of fairly uniform "high" quality. When many choice criteria are applied to information which represents brands to be of relatively lower quality (but very good on some attributes and very poor on others), the participants responded by shrinking their evoked set size. This suggests at least two observations.

First, that people appear to construct an evoked set of a size dependent on the conditions of the study implies that greater attention should be paid to this phenomenon in information processing studies. Since there appears to be no way to directly manipulate the size of an evoked set, it can operate as an uncontrolled and possible confounding variable in information processing experiments. Although there does not appear to be any evidence bearing on the issue, it seems reasonable to assume that "overall" ratings (or rankings) of brands which are not included in the evoked set (i.e., are in the "inept" or "inert" sets) would be less reliable and/or less valid than those of brands included in the evoked set. Studies which attempt to fit models to "overall evaluation" ratings (or rankings) for all brands in the awareness set (or about which information is provided in an experimental setting) should probably closely consider the saliency of the attributes presented and the variability and mean levels of the information to which participants are exposed.

Second, it has been argued that, for persons using more choice criteria, decisions are more complex (Park, 1976). This seems reasonable as long as decision makers are combining attribute information in a manner simulated by one of the compensatory models. Indeed, the finding that persons using more choice criteria constructed smaller evoked sets may be explained by arguing that they anticipated the large number of comparisons which would have to be made and, thus, felt strongly impelled to reduce their evoked set size. However, as noted by Reilly and Holman (1977), using the "number of choice criteria" as a stratifying variable creates self-selected treatment groups and causal inferences become impossible. On the other hand, because there are substantial inter-individual differences in the number and saliency of choice criteria, manipulation of this variable is, if not impossible, so difficult as to rule out a straight-forward test of the relationship between the number of criteria used and complexity. In the present

study, since the evoked set shrank as the variability of attribute information increased (while the mean levels decreased), it appears more parsimonious to infer that the use of more choice criteria increased the likelihood of "knocking brands out" by finding some attribute which failed to achieve a minimum acceptance level. To the extent that persons "sort" brands to construct an evoked set, it appears that the use of more choice criteria does not increase complexity but, rather, facilitates simplification.

References

Mark I. Alpert, "Identification of Determinant Attributes: A Comparison of Methods," Journal of Marketing Research, 8 (May, 1971), 184-91.

James R. Bettman, "Data Collection and Analysis Approaches for Studying Consumer Information Processing," in W. D. Perreault, Jr., ed., Advances in Consumer Research, Vol. 4, Proceedings of the 7th Annual Conference of the Association for Consumer Research, 1977, 342-48.

Brian M. Campbell, "The Existence of Evoked Set and Determinants of Its Magnitude in Brand Choice," Unpublished doctoral dissertation, Columbia University, 1969.

Steven H. Chaffee and Jack M. McLeod, "Consumer Decisions and Information Use," in S. Ward and T. S. Robertson, eds., Consumer Behavior: Theoretical Sources, Englewood Cliffs, N. J.: Prentice-Hall, 1973, 385-415.

Jeremy D. Finn, Multivariate: Univariate and Multivariate Analysis of Variance, Covariance and Regression, (Chicago: National Educational Service, 1972).

Kjelf Gronhaug, "Some Factors Influencing the Size of the Buyers Evoked Set," European Journal of Marketing, 7 (Winter 1973/74), 232-41.

John A. Howard and Jagdish N. Sheth, The Theory of Buyer Behavior, (New York: Wiley, 1969).

Lance P. Jarvis and James B. Wilcox, "Evoked Set Size -- Some Theoretical Foundations and Empirical Evidence," in T. V. Greer, ed., Combined Proceedings Fall Conference of the American Marketing Association, No. 35, 1973, 236-40.

Roger E. Kirk, Experimental Design: Procedures for the Behavioral Sciences, (Belmont, Calif.: Brooks/Cole, 1968).

Frederick E. May and Richard E. Homans, "Evoked Set Size and the Level of Information Processing in Product Comprehension and Choice Criteria," in W. D. Perreault, Jr., ed., Advances in Consumer Research, Vol. 4, Proceedings of the 7th Annual Conference of the Association for Consumer Research, 1977, 172-75.

James H. Myers and Mark I. Alpert, "Semantic Confusion in Attitude Research: Saliency vs. Importance vs. Determinance," in W. D. Perreault, Jr., ed., Advances in Consumer Research, Vol. 4, Proceedings of the 7th Annual Conference of the Association for Consumer Research, 1977, 106-10.

Chem L. Narayana and Rom J. Markin, "Consumer Behavior and Product Performance: An Alternative Conceptualization," Journal of Marketing, 39 (October, 1975), 1-6.

Lyman E. Ostlund, "Evoked Set Size: Some Empirical Results," in T. V. Greer, ed., Combined Proceedings Fall Conference of the American Marketing Association, No. 35, 1973, 226-30.

Whan C. Park, "The Effect of Individual and Situation-Related Factors on Consumer Selection of Judgmental Models," Journal of Marketing Research, 13 (May, 1976), 144-51.

Bernard Pras and John Summers, "A Comparison of Linear and Nonlinear Evaluation Process Models," Journal of Marketing Research, 12 (August, 1975), 276-81.

Michael Reilly and Rebecca H. Holman, "Does Task Complexity of Cue Interrelation Affect Choice of an Information Processing Strategy? An Empirical Investigation," in W. D. Perreault, Jr., ed., Advances in Consumer Research, Vol. 4, Proceedings of the 7th Annual Conference of the Association for Consumer Research, 1977, 185-90.

CONSUMER-ORIENTED VERSUS ADVERTISER-ORIENTED
LANGUAGE: COMPREHENSIBILITY AND SALIENCE OF THE
ADVERTISING MESSAGE

Martin R. Lautman, Associates for Research in Behavior, Inc.
Larry Percy, Gardner Advertising, Co.

Abstract

While the effectiveness of advertising is directly related to a receiver's ability to comprehend its message, too frequently advertising utilizes language which is neither salient nor understandable to the receiver. This paper addresses this problem of advertiser-consumer language compatibility in advertising, and illustrates a method, based upon a qualitative-quantitative research progression, for insuring understandable and meaningful copy in one's advertising.

Consumer-oriented Versus Advertiser-oriented Language:
Comprehensibility and Salience of the
Advertising Message

The effectiveness of an advertisement is directly related to the ability of a target audience to comprehend its message. This comprehension is a function of the message being presented in a manner highly compatible in both linguistic structure and semantic content with that commonly experienced and expected by the target audience (Lautman, 1974). In the broadest sense, to the extent the audience must make a series of simplifying "transformations" of the message to understand it, comprehension can be expected to suffer. Various studies have been directed toward these issues. For example, it has been shown that certain sentence structures (Wright, 1969; Olson and Filby, 1972) and various technical wordings (Lautman, Siegel, and Burkett, 1973; Siegel, Lautman, and Burkett, 1974) tend to affect the comprehensibility of a prose message.

The results of these and other psycholinguistic studies directly relate to the issue of advertising comprehensibility. Frequently, an advertiser is concerned that detailed or technical attributes of his product be clearly communicated to consumers. He is apt to select some specific technological superiority and use that "benefit" as the basis for differentiating his product. Or, more commonly, he may wish to include all of the features he perceives to be important, regardless of their inherent meaningfulness or salience to the consumer.

In attempting to execute copy describing these technical features, a copywriter often finds that descriptions and terminology must be included which may have little meaning to the consumer. More often than not, there is a strong tendency to use the advertiser's technical terminology rather than the consumer's language in composing these product descriptions.

A somewhat different scenario with similar results occurs when the advertiser uses terminology which over time has acquired alternative meanings. In the pharmaceutical area, for example, usage of the term "broader spectrum" for an antibiotic can mean either effective against more strains of a single type of bacteria or effective against more types of bacteria. Once again, in order to accurately advertise his product and make clear its advantages the advertiser is faced with understanding how his target audience perceives this frequently used technical term.

It is with these concerns in mind that this study is directed towards the problem of advertiser-consumer language compatibility and illustrates a method for insuring understandable and meaningful copy in one's advertising.

Key to this approach is a research strategy based on a qualitative-quantitative progression. Qualitative exploratory work alone, while generally able to develop a consumer language and identify salient attributes, may fail to answer for the advertiser both how best to use the consumer's language in his advertising, as well as how the consumer comprehends and perceives the salience of competitive messages. Exploratory quantitative work, while generally able to deal with these latter two questions effectively, without the benefit of the earlier qualitative exploration runs the dual risk of not identifying the most salient areas to the consumer and not developing an advertising language meaningful to him.

It should be pointed out that this qualitative-quantitative progression should not be viewed in the usual "hypotheses generating-validating" mode frequently assigned the step-wise use of focus groups and subsequent survey work. Rather, each contributes in its own right, with the initial qualitative work generating the parameters of consumer language and the subsequent quantitative exploration comparing those parameters with existing advertiser language.

This distinction is in part discussed by Calder (1976) in describing approaches for evaluating qualitative research goals. His "exploratory" approach follows the typical pattern of using focus groups to generate or select hypotheses which one plans to verify by fielding a large sample survey. A second approach, referred to as "clinical," belies the need for quantitative follow-up and relies on the "clinical" judgment of a specially trained analyst--frequently a psychologist--to reveal the underlying causes of behavior. His third approach closely follows the role of qualitative research as it is considered here within the qualitative-quantitative progression as applied to consumer vs. advertiser language. This Calder terms the "phenomenological" approach.

The phenomenological approach to qualitative research offers a description of how consumers perceive reality in their own terms. Calder points out that the goals of what he calls the phenomenological approach to qualitative marketing research is identical to that of phenomenological sociology as derived from Schutz (1967). For this reason, the phenomenological approach is ideal for describing the role qualitative research must play in the progression being suggested here for the study of consumer-advertiser language compatibility.

The advertiser for the most part, belongs to a social grouping whose perceptions are generally not the same as their target market. Reality to the advertiser is frequently divorced from that of the consumer. This appears to be especially true when the advertised products tend to be more "technologically" oriented (for example, electronics equipment) or when common usage of technical terms among a group of specialists (such

as physicians) tends to vary from traditional definitions. Using qualitative research to understand how consumers perceive a product, in their own terms, helps to bridge this social distance between the advertiser and consumer.

Quantitative research, still in an exploratory mode, is then able to juxtapose the consumer developed language from the focus groups with that used by the advertiser in an effort to sort out commonalities and misinterpretations, and provide guidance for those message points, again based upon consumer language, important to a positive buyer response to the advertising.

Study Design

In an effort to develop a consumer-relevant language for an electronics product, and assess the degree of congruence between the consumer-relevant language and that used by various advertisers in the category in describing product benefits, a two-part research study following the qualitative-quantitative progression outlined above was conducted. A general outline of the study design is shown in Figure 1.

FIGURE 1
QUALITATIVE/QUANTITATIVE PROGRESSION

Focus group discussions using a phenomenological approach to determine the parameters of consumer language and the salient attributes.

Exploratory quantitative research using personal interviews to evaluate (1) the comprehension of existing advertiser language by consumers, and (2) belief congruence and preferences within and between advertiser appeals and consumer desired attributes.

Method

Six focus groups were conducted among men and women owners, and those expressing an interest in owning an electronic product. Three groups were conducted in St. Louis and three in Philadelphia. In addition to the global concern of generating the parameters of consumer language, information was sought concerning awareness of brands in the category, specific evaluations of different product features, situational effects, etc. as they related to consumer conceptualizations of the entire product area.

Follow-up exploratory quantitative work was conducted among 120 male and female respondents in three cities: Pittsburgh, Tampa, and Phoenix. Personal interviews were conducted with a sample selected on an area-wide cluster basis, stratified to insure representation along a broad demographic base. Two basic exercises were conducted: 1) pairwise similarity data and a preference ordering in terms of intention to buy were gathered among selected advertising claims; 2) perceived pairwise belief congruence strength was measured among consumer suggested features along with a preference ordering in terms of importance of the features in a brand.

In the first exercise each subject was asked to provide pairwise similarity data on eight brand appeals currently in use by major category manufacturers by sorting a set of 28 cards, where each card contained one of the $n(n-1)/2$ pairs of appeals. Subjects were asked to look through all of the cards, placing those cards they felt contained similar appeals into one pile and all those cards they felt contained appeals which were not similar into a second. They were then asked to rank the first pile from the card containing the two most

similar appeals, and to rank the cards in the second pile from those containing the two most dissimilar appeals. The order of the second pile was then inverted and added to the first, providing a ranking or all 28 pairs from most to least alike. Following this sort, each subject ranked the appeals in ascending order according to their likelihood of buying a product if all they knew about that particular brand was that it was advertised in the way described on each card.

Next, pairwise belief congruence ratings were collected for each of 28 pairs of salient attributes described in consumer-relevant terms, as identified by consumers in the initial qualitative focus groups (each pair representing one of the $n(n-1)/2$ pairs of eight attributes). Using a five-point Likert-type belief scale, each respondent was asked whether they thought a typical brand in this category that contained one attribute feature would also contain another.

TABLE 1
ADVERTISER APPEALS AND CONSUMER-DERIVED ATTRIBUTES

Advertiser-Appeals	Consumer-Derived Attributes
1. Dependability	1. Solid guarantee
2. Nurturant I	2. Powerful
3. Personification	3. Rugged
4. Nurturant II	4. Easily repaired
5. Power	5. Moderately priced
6. Adventure	6. Solid state
7. Strength	7. Portable
8. Rugged	8. Foreign made

Finally, considering each attribute individually, they were asked to rank the features from most to least important in terms of their brand choice decision. The eight advertiser-derived appeals and eight consumer-derived attributes are shown in Table 1.

Analytic Design

The data collected in the pairwise similarity rankings of the eight current advertiser-derived appeals were submitted to Kruskal's (1969) M-D-SCAL 5M program for non-metric multidimensional scaling. Solutions were sought in three, two and one dimensions. The rank-order "likelihood-to-buy" data were then imbedded into the solution space via Carroll and Chang's (1967) PREFMAP. Solutions were explored for both Phases III and IV. This joint solution provides a geometric interpretation of the perceived cognitive similarity among existing product appeals as expressed in advertiser language, along with "ideal points" reflecting those appeals most likely to attract consumer interest.

The second exercise provided mean ratings on a five-point Likert-type belief scale for each pair of eight consumer-derived attributes. Rank ordering these means provided a belief proximity measure which was introduced into Kruskal's (1969) M-D-SCAL 5M program, solving for three, two, and one dimensional solutions. The rank order preference data for the eight attributes were then imbedded into the solution space via PREFMAP giving a geometrical representation of the belief congruence associated with the consumer-derived attributes and the salience of those attributes with respect to product preference.

Results

The first step in our qualitative-quantitative progression was to generate an understanding of how consumers talk about specific consumer electronics products. The results of the focus group discussions revealed a set of eight specific product attributes, described in the consumer's own language. In addition, probing of spe-

cific technical terms indicated little consumer understanding of what a manufacturer meant when using those terms. These findings provided the necessary consumer input for the exploratory quantitative follow-up, along with a strong indication that the advertiser's language in this category was generally not understood by the consumer.

The next step was to utilize these results, along with a set of advertiser-derived appeals, in an effort to determine the impact of each on general consumer understanding and preference. Utilizing Kruskal's M-D-SCAL 5M to analyze the $n(n-1)/2$ pairwise similarities rankings of the eight current advertiser-derived appeals considered provided a usable solution configuration in two dimensions. The results shown in Figure 2 demonstrate a significant amount of perceived discrimination among the various advertiser-derived appeals. It would appear that the executions of the Dependability, Rugged, and Strength appeals, which one might reasonably expect to be quite similar, are in fact perceived as highly dissimilar. While it should be cautioned that other executions of these same appeals could produce different results, it is nonetheless significant that the appeals of these three advertisers, as expressed in their language, were not perceived to have much in common by the consumer.

FIGURE 2
NON-METRIC MULTIDIMENSIONAL SCALING CONFIGURATION
FOR EIGHT ADVERTISER-DERIVED APPEALS FROM
M-D-SCAL 5M

Pairwise Similarity
Rugged Appeal

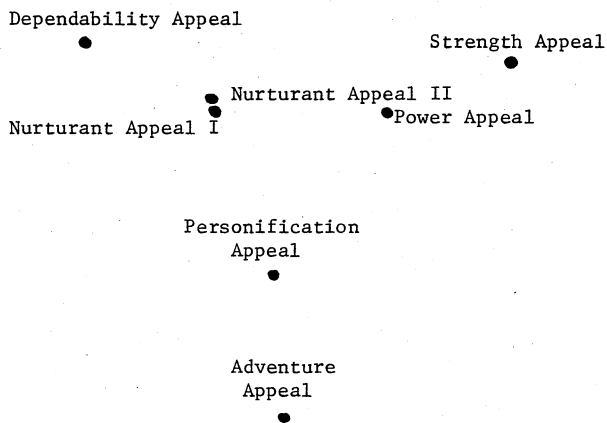
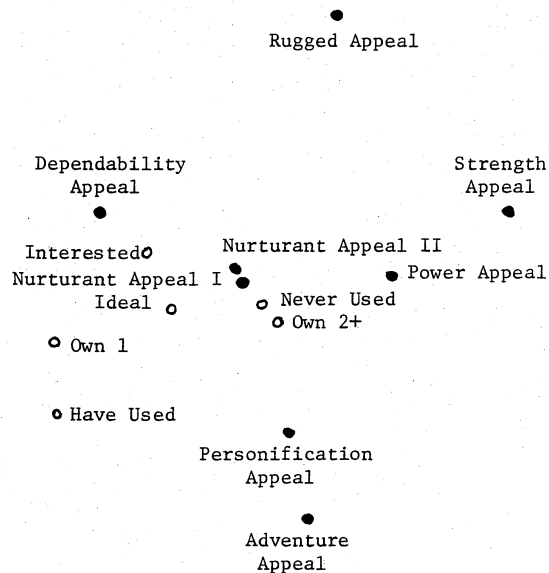


Figure 3 shows the results of a PREFMAP Phase III solution based on the likelihood to purchase given the advertiser-derived appeals. Those appeals most likely to stimulate consumer interest tend to be the Dependability appeal and the two Nurturant appeals.

FIGURE 3
TWO-DIMENSIONAL JOINT PLOT OF ADVERTISER-DERIVED
APPEALS WITH PREFERRED APPEAL FROM PREFMAP
PHASE III ANALYSIS



In an attempt to judge the extent of belief congruence among the consumer-derived attributes, the beliefs for each of the $n(n-1)/2$ pairs of attributes were analyzed via Kruskal's M-D-SCAL 5M. The resulting configuration shown in Figure 4 suggests that while consumers believed a product may be rugged, solid state, and have a solid guarantee, a product with these attributes had little in common with a product described by any of the other attributes. A powerful product, for example, was unlikely to be believed easily repaired or foreign made.

FIGURE 4
NON-METRIC MULTIDIMENSIONAL SCALING CONFIGURATION
FOR EIGHT CONSUMER-DERIVED ATTRIBUTE BELIEFS
FROM M-D-SCAL 5M

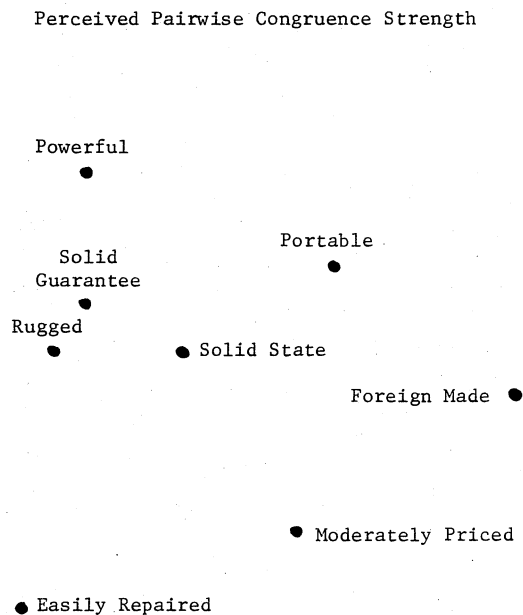
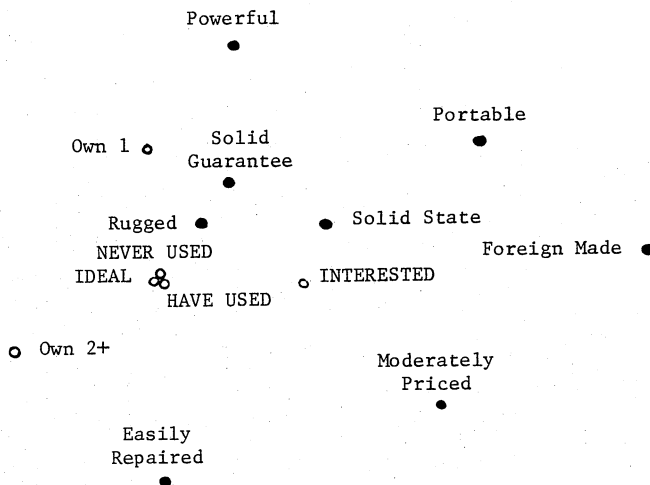


Figure 5 shows the results of a PREFMAP Phase III analysis of preference within the belief congruence solution configuration. It shows that while the rugged-solid guarantee features tended to be most preferred among non-owners, single product owners skewed towards powerful, and owners of two or more towards easily repaired.

Perhaps the most interesting overall finding was the extent to which little compatibility seemed to exist between the advertiser-derived appeals and the consumer-derived attributes. Advertisers were not generally addressing consumer wants with consumer language. Only the Dependability appeal appeared to be compatible with the "rugged - solid guarantee - solid state" cluster of consumer-described attributes. The advertiser's version of Rugged and Strength were not perceived as such by the consumer; and the advertiser's interest in Nurturant appeals (and miscellaneous others) did not seem to be reflected in desired consumer attributes.

FIGURE 5
TWO-DIMENSIONAL JOINT PLOT OF CONSUMER-DERIVED
ATTRIBUTE BELIEFS WITH PREFERRED ATTRIBUTE
ASSOCIATIONS FROM PREFMAP PHASE III ANALYSIS



The technical approach outlined here did not include a direct joint comparison of advertiser appeals and consumer-derived attributes. Two separate analyses were conducted and the results were compared. A more direct approach, and one which is advocated, is to have respondents compare the appeals in terms of the attributes. For example, the appeals might be rank ordered or rated in terms of their perceived offering of each salient attribute. An unfolding analysis might then show how much each attribute is associated with each appeal. Appeals viewed as best in terms of offering positively salient attributes might then be selected. When the observed incompatibilities are not as great as they were here, this direct approach will be a more sensitive test.

Discussion and Conclusion

While much research has been directed towards pinpointing consumer needs and preferences (for example, Haley, 1968; Sweitzer, 1975) and selecting advertising themes and executions which reflect them (for example, Percy, Lautman, and Kordish, 1976; Smith and Lusch, 1976) little, if any, research has been channeled towards directly studying the advertiser's use of language in the copy

of the ads themselves. This study has shown that for at least one technical product marketed to a mass audience, serious communication problems may exist. It was demonstrated that consumers (1) often do not understand the advertiser's use of technical language in promoting the product and (2) do not necessarily interpret advertising appeals designed to promote a particular product benefit or benefits, as was intended, finding the advertiser's language incompatible with their own consumer-oriented needs. The qualitative-quantitative research progression outlined and demonstrated here provides a methodology for identifying and bridging language gaps between the advertiser and consumer, and helping to assure comprehensible advertising appeals describing positively salient product benefits.

It is important to recognize that the results of this study should not be interpreted as suggesting that all technical wording be removed from advertising. Obviously, in the pharmaceutical, industrial and other areas advertising must be technical. What is being recommended, however, is that the advertiser test his use of technical terminology so as to insure that his target market both understands it and views it in the context of his appeal as directed towards promoting positively salient attributes.

Compatibility problems between advertiser and consumer-oriented appeals emanating from language can have various sources. First, particular words may not be generally understood. Second, particular words while generally understood may have multiple meanings and the context in which they are presented may not clearly identify which one is appropriate. Third, certain syntactic structures are more difficult to process and comprehend. Finally, the theme of the advertising may be misinterpreted or missed entirely through problems such as poor copy organization, inappropriate use of examples, unexpected or unusual presentations of the copy, etc. This study dealt only with the first, second, and, to a certain extent, the last of these sources of problems. Research on the third problem, the comprehensibility of various sentence structures, while studied extensively by psycholinguists (see Wright, 1972; Garrod and Trabasso, 1973; Kanouse, 1972; Abelson and Kanouse, 1966), has not often been directly investigated by consumer researchers in terms of identifying prose structures most suitable for advertising communication. The authors have found in an as yet unreported study that variations in sentence structure can have a significant effect on advertising headline believability, suggesting more inquiry into this area by researchers interested in consumer response to communications.

Perhaps the most important conclusion which can be drawn from this study is that advertisers cannot take for granted that a target audience will comprehend and interpret their appeals/language in the manner intended by the agency (Lautman, Percy, and Kordish, in press). In the process of translating advertising objectives into copy, care must be taken that the creative approach should not overwhelm or distort the intended communication. Even when technical terminology is not involved, usage of a consumer-oriented language can be expected to lead to easier comprehensibility of the advertiser's message. Testing that the advertiser's intended message is, in fact, the one being communicated to the target audience is a necessary step in assessing the effectiveness of any advertising copy.

References

- R. P. Abelson and D. E. Kanouse, "Subjective Acceptance of Verbal Generalization," In S. Feldman (Ed.), Cognitive Consistency: Motivational Antecedents and Behavioral Consequents, New York, Academic Press, 1966.

B. J. Calder, "Focus Groups and the Nature of Qualitative Marketing Research," Working Paper No. 76-23, Northwestern University, August, 1976.

J. D. Carroll and J. J. Chang, "Relating Preference Data to Multidimensional Scaling Solutions via a Generalization of Coomb's Unfolding Model," Bell Telephone Laboratories, Murry Hill, N. J., 1967.

S. Garrod and T. Trabasso, "A Dual-Memory Information Processing Interpretation of Sentence Comprehension," Journal of Verbal Learning and Verbal Behavior, 12, 155-167.

R. I. Haley, "Benefit Segmentation: A Decision-Oriented Research Tool," 32 (July, 1968), 30-35.

D. E. Kanouse, "Language, Labeling, and Attribution," In E. E. Jones, et al. (Eds), Attribution: Perceiving the Causes of Behavior, Morristown, N. J., General Learning Press.

J. B. Kruskal and F. J. Carmone, "How to Use M-D-SCAL, A Program to do Multidimensional Scaling and Multidimensional Unfolding," (Version 5M of M-D-SCAL, all in Fortran IV), Bell Telephone Laboratories, Murry Hill, N. J., 1969.

M. R. Lautman, "Psycholinguistics and Consumer Awareness: The Message is the Medium," Paper presented at the 82nd Annual American Psychological Association Convention, New Orleans, Louisiana, 1974.

M. R. Lautman, L. Percy and G. Kordish, "Developing and Testing Advertising Campaigns Using Multidimensional Scaling Techniques," Journal of Advertising Research (in press).

M. R. Lautman, A. I. Siegel, and R. Burkett, "Effects of Simplification and Multimodal Presentation on the Comprehensibility of Technical Training Materials," Proceedings of the 81st Annual American Psychological Association, 8, 1973, 737-738.

D. R. Olson and N. Filby, "On the Comprehension of Active and Passive Sentences," Cognitive Psychology, 3, 1972, 361-381.

L. Percy, M. R. Lautman, and G. R. Kordish, "A Non-Metric Approach for Developing and Evaluating Alternative Advertising Strategies," Advances in Consumer Research IV, 1976, 191-196.

A. Schutz, The Phenomenology of the Social World, Evanston, Illinois, Northwestern University Press, 1967.

A. I. Siegel, M. R. Lautman, and R. Burkett, "Reading Grade Level Adjustment and Auditory Supplementation as Techniques for Increasing Textual Comprehensibility," Journal of Educational Psychology, 66, 1974, 895-902.

R. E. Smith and R. F. Lusch, "How Advertising Can Position a Brand," Journal of Advertising Research, 16 (February, 1976), 37-43.

R. W. Sweitzer, "Determinant Attribute Segmentation," Journal of the Academy of Marketing Science, 3 (Winter, 1975), 85-98.

P. Wright, "Two Studies of the Depth Hypothesis," British Journal of Psychology, 60, 1969, 63-69.

_____, "Some Observations on How People Answer Questions About Sentences," Journal of Verbal Learning and Verbal Behavior, 11, 1972, 188-195.

WHO IS THE DEAL PRONE CONSUMER?¹

Robert Blattberg, University of Chicago
Thomas Buesing (student), University of Chicago
Peter Peacock, Wake Forest University
Subrata Sen, University of Rochester

Abstract

This paper uses a model of consumer buying behavior to identify household characteristics that should affect deal proneness. The model assumes that the household makes purchasing and inventory decisions much like a firm. In other words, the household's purchasing decisions are based on factors such as transaction costs, holding costs, and stockout costs in addition to product price. Household characteristics are then related to these cost parameters to predict which households are likely to be deal prone. The predictions are tested empirically using panel data on five frequently-purchased products. The empirical results indicate that it is possible to identify the deal prone household and that the key variables that affect deal proneness are household resource variables such as home ownership and automobile ownership.

Introduction

Marketing managers have always been interested in identifying the deal-prone household on the basis of available demographic data. If such households can be identified precisely, specific marketing strategies designed to appeal to such households are likely to be more effective. For example, demographic information is available by zip codes or census tracts. If certain demographic groups are more deal prone, coupon distribution could be restricted to those areas where households with higher deal proneness reside. This would reduce couponing costs with a less than proportionate reduction in response. Similarly, more accurate identification of deal prone households would increase the marketer's ability to match deal prone households and media audience characteristics, thus increasing the efficiency of media distribution of coupons and other promotional items.

Several studies have tried to identify the deal prone household. Webster (1965) and Montgomery (1971) have published two of the better known studies. The results of these and other studies are summarized by Frank, Massy, and Wind (1972, p. 124) who state:

The results of cross-sectional studies, almost without exception, indicate that there is, at best, only a modest degree of association between demographic, socioeconomic, and/or personality characteristics, and selected aspects of household purchasing behavior, such as total consumption, brand loyalty, and deal proneness.

One reason for this "modest degree of association" may lie in the methodological approach usually taken in these studies. Typically, a large number of potential explanatory variables are regressed against the proportion of purchases made on deal in a search for statistical significance. For example, Webster (1965) ran 200 regressions with different combinations of 45 explanatory variables. This approach is open to serious question because one cannot always determine if "signi-

ficant" relationships reflect a valid relationship or a spurious one which has arisen by chance alone. Without a theory to indicate which variables should affect deal proneness, there is a greater danger of accepting spurious results.

A related deficiency of prior studies of deal proneness arising from the absence of a clearly stated theory has been improper specification of explanatory variables. For example, Montgomery (1971) included "presence of children" as an independent variable in a model but was unable to predict a priori whether the presence of children should or should not increase deal proneness. However, if one hypothesizes that it is the age of the children that affects deal proneness (rather than their presence per se), it is possible to predict a priori the impact of "age of children" upon deal proneness. If the children are below the age of six (and are consequently not yet in school), they require more of their parents' time, thus reducing the time available for shopping. Having less time to shop leads to fewer shopping trips and fewer opportunities to take advantage of deals, thereby reducing the household's deal proneness. Montgomery's results led him to conclude that presence of children was not related to deal proneness. However, by considering the presence of children without determining how their presence should affect deal proneness, Montgomery may have arrived at an incorrect conclusion about a variable which, if properly specified, could well be related to deal proneness.

Like the studies cited above, this paper also attempts to identify the deal prone household. However, the approach used here is quite different from that used in most of the earlier work. We begin by formulating a model of household purchasing behavior. The model is then used to predict how certain demographic variables should affect deal proneness. Finally, an empirical evaluation of the predictions is made. The empirical results show that it is possible to identify the deal prone household using demographic variables and that the effect of these variables is substantial.

The Model of Household Purchasing Behavior

Model Assumptions

A household inventory model is developed in this section which is based on the assumption that the household is the same as any producing unit which needs to stock inventory and meet demand. This assumption follows from the notion of the household as a production unit which Becker (1965) and others have used to model consumer behavior in the economics literature. This modeling approach is used because it has proved to be very fruitful in several applied studies in economics (Michael and Becker, 1973) and also because it appears to be a potentially promising approach in marketing (Kunreuther, 1973). The inventory model proposed here also corresponds closely to models developed by management scientists to make better inventory decisions in more traditional production environments (Hillier and Lieberman, 1974, pp. 472-527).

¹This research has been funded, in part, by National Science Foundation Grant SOC73-05547.

The initial assumption in the model is that households make long-run decisions about whether to use a given product at all and the average number of units of the product to use per period. Decisions about product use and the average usage rate are determined exogenously by factors such as family size, family income, etc. It is also assumed that ratios of prices of the product in question and prices of substitutes and complements are constant during the period considered so that households need not evaluate their usage rate decision because of changes in relative price. The latter assumption is made in order to obtain a tractable model and because it permits us to concentrate on the purchase timing decision which is the basic focus of this paper.

Cost Structure of the Household

Four categories of cost affect household inventory decisions: (1) transaction cost, (2) storage cost, (3) stockout cost, and (4) the actual price of the item. Transaction cost is the opportunity cost of the time required to purchase an item once the consumer is actually in a store plus the opportunity cost of travel time required to get to and from the store where the purchase takes place. Transaction cost will vary across stores. And if a consumer has a "regular" or preferred store, one would expect the transaction cost of an item purchased there to be less than if a special trip is made to purchase the item at some other store. Storage cost represents interest on the capital required to maintain a given level of inventory plus the cost of the required space. Stockout cost reflects the foregone utility of not consuming an item which is not in stock at the time it is demanded. If the household can easily substitute other items in the event of stockout, or if it derives little utility from consuming the item, stockout cost should be low. Observed price per unit is the final component of cost. For purposes of the analysis, it is assumed that the observed price in a store is constant within any given period, e.g., a week. Prices may vary across stores and may change from period to period.

Mathematical Formulation of the Model

The household's purchase decision process is represented mathematically in Exhibit 1. The model shown requires the consumer to minimize expected costs over a finite time horizon which includes present as well as future periods. Thus, expectations about future prices and future demand affect the present period's decision. Note that though the household's average demand per period is known, the exact quantity demanded in each period is unknown at the start of the period. Thus, this quantity, \tilde{d}_t , is represented as a random variable in the model. Future price expectations are generated by probability distributions of the time between deals and the length of time a given deal is in effect. These distributions would be based on actual household experience with deal duration and time between deals in each store.

For the model shown in Exhibit 1, it is assumed implicitly that all brands in the product class yield the same utility to the consumer. Therefore, the consumer's objective is simply to minimize the "total cost" of buying the product. Again, this assumption is made only to obtain a tractable model. In order to consider the case where different brands offer different utilities, a framework which jointly considers utility maximization and purchase timing would have to be developed.

EXHIBIT 1 MATHEMATICAL FORMULATION OF THE MODEL OF HOUSEHOLD PURCHASING BEHAVIOR

$$\min_{X_{i,t}} E \left\{ \sum_t \sum_i f_{i,t} [X_{i,t}] + h_t \tilde{I}_{t-1} + u_t \tilde{S}_t \right\}$$

subject to:

$$f_{i,t} [X_{i,t}] = \begin{cases} T_{i,t} + P_{i,t} X_{i,t} & \text{if } X_{i,t} > 0 \text{ for } i=1, \dots, K \\ 0 & \text{if } X_{i,t} = 0 \end{cases}$$

$$\tilde{I}_t = \begin{cases} I_{t-1} + \sum_i X_{i,t} - \tilde{d}_t & \text{if } \tilde{d}_t < I_{t-1} + \sum_i X_{i,t} \\ 0 & \text{if } \tilde{d}_t \geq I_{t-1} + \sum_i X_{i,t} \end{cases}$$

$$\tilde{S}_t = \begin{cases} \tilde{d}_t - \tilde{I}_{t-1} - \sum_i X_{i,t} & \text{if } \tilde{d}_t > I_{t-1} + \sum_i X_{i,t} \\ 0 & \text{if } \tilde{d}_t \leq I_{t-1} + \sum_i X_{i,t} \end{cases}$$

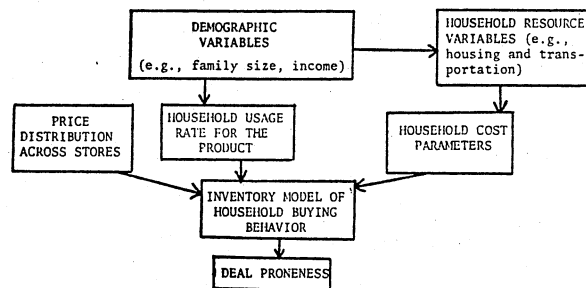
where:

- $T_{i,t}$ = transaction cost at store i at time t ,
- $P_{i,t}$ = price per unit of product at store i at time t ,
- $X_{i,t}$ = quantity purchased at store i at time t ,
- I_t = inventory on hand at beginning of time t ,
- \tilde{d}_t = quantity demanded at time t ,
- h_t = unit holding cost at time t ,
- \tilde{S}_t = amount of stockout at time t ,
- u_t = unit stockout cost at time t ,
- K = the number of stores.

Value of the Inventory Model

The inventory model described above provides a theoretical basis to identify demographic and household resource variables which should lead to deal proneness. To see this, consider Exhibit 2 which provides a diagrammatic representation of the manner in which the inventory model links demographic and resource variables with household deal proneness. Exhibit 2 indicates that the

EXHIBIT 2 DEAL PRONENESS AND HOUSEHOLD DEMOGRAPHIC AND RESOURCE VARIABLES



three inputs to the inventory model are: (1) the price distribution of the product across the stores, (2) the household usage rate for the product (which, in turn, is influenced by demographic variables such as family size) and (3) household cost parameters.

Deal Proneness and Household Cost Parameters

The first link of interest is the one that connects household cost parameters with deal proneness through

the inventory model. What does this link represent? Essentially, it indicates how the cost structure facing a household determines whether or not the household will be deal prone. For example, if storage costs were low, one would expect households to stock up on a commodity when a deal is on. Alternatively, if the transaction cost of traveling to a non-preferred store were sufficiently high, one would not anticipate a consumer to buy on a deal offered by the non-preferred store.

Cost Parameters and Household Demographics and Resources

But what factors determine the household's cost structure? To answer this question, consider the link in Exhibit 2 which connects Demographic Variables (like income) to Household Cost Parameters through Household Resource Variables (such as housing and transportation). Note first that income is an important determinant of household resources, i.e., households with higher income are more likely to own homes (as opposed to being renters) and are also more likely to own one or more cars. These household resources, in turn, affect the cost parameters of the model. For example, home owners typically have more storage space available compared to apartment dwellers and hence should incur lower storage costs. Similarly, car ownership makes transportation easier, thereby reducing the household's transaction costs.

It was noted earlier that low storage costs and low transaction costs both lead to deal proneness. Since low storage costs are associated with home ownership and low transaction costs with car ownership, it is possible to make specific predictions such as the following: home owners and car owners will tend to be more deal prone than apartment dwellers and households without cars.

One can see, therefore, that the inventory model's links with deal proneness and with household cost parameters makes it possible to (1) identify the relevant variables that should affect deal proneness and (2) predict the direction of their effect. The model is also useful in providing a quantitative evaluation of the relative impact of the various cost parameters on deal proneness (see Blattberg, et. al., 1977).

The Effect of Demographics and Household Resource Variables on Deal Proneness

In this section, we state formally our predictions of how some specific household resource and demographic variables lead to deal proneness. Because of the length constraints on this paper, only two types of explanatory variables will be studied: (1) the household resource variables mentioned earlier, i.e., car ownership and home ownership, and (2) income. Note that the data for these variables are available by zip codes or census tracts. Therefore, if these variables do affect deal proneness, the marketing manager can implement the results easily. This is in contrast to some of the variables found by Webster (1965) and Montgomery (1971) to affect deal proneness. For example, both found that brand loyalty was negatively associated with deal proneness. However, one must first identify who the less brand loyal consumers are before one can use such a finding.

Household Resource Variables and Deal Proneness

A key component of the transaction cost of shopping is transportation cost. Households that do not have cars available are forced to shop at stores that are nearby. They are also more likely to shop at a single store (see Kunreuther, 1973, p. 376). Since the ability to take advantage of deals depends upon the freedom to shop often and at many stores, households without cars should be less deal prone.

The second household resource variable is home ownership. This variable should be related to holding costs. Apartment dwellers usually have less storage space available than homeowners simply because apartments are smaller. Therefore, holding costs should be higher for apartment dwellers. Since lower holding costs should lead to greater deal proneness, homeowners should be more deal prone than apartment dwellers.

The Effect of Income on Deal Proneness

The usual argument given in support of a negative relationship between deal proneness and income is that low income households have lower time costs resulting in lower search and transaction costs. Furthermore, economic theory suggests that lower income households should be more price sensitive. Empirical research in marketing has rarely shown that income affects deal proneness (see Webster, 1965, for example). If an effect is found at all, higher income seems to lead to greater deal proneness rather than less.

The problem with studying the effects of income is that income effects are confounded by the effects of household resource variables. The influence of these other variables must be removed before the effect of income can be clearly evaluated. It was stated above that car and home ownership should increase deal proneness. Higher income households are more likely to buy capital goods such as cars and homes, thus increasing their deal proneness. The resulting interaction between income and household resources may result in the anomalous finding that high income households are more deal prone than low income households. If resources available were held constant, however, one should observe the opposite outcome.

Empirical Results

The data used to analyze deal proneness were the Chicago Tribune Panel purchase data and associated demographic variables. Consumers classified into three segments defined by Blattberg and Sen (1976): the National Brand Loyal Deal, the National Brand Switcher Deal, and Deal-Oriented, are defined here as being deal-oriented. All consumers classified into one of the other stable pattern categories (i.e., not including the Changing Pattern or Last Purchase Loyal patterns) constituted the non-deal prone population.¹ Five product categories were studied: aluminum foil, waxed paper, headache remedies, liquid detergent and facial tissue. The data were from 1958 to 1966, depending on the category.² The household variables studied are those described in the previous section.

Blattberg and Sen (1976) classified each household's purchase patterns into segments. Deal proneness was based on membership in one of the above three segments and was a dichotomous variable: deal prone or not deal prone. Given that the effect of the independent variables on deal proneness may be nonlinear, it was decided to use cross-classification analysis instead of regression (see Frank, Massy and Wind, 1972, pp. 126-129). A major problem in doing cross-classification analysis

¹The nonstable buying patterns contain some deal prone households who for certain periods of the data were deal prone and for other periods were not. They have been excluded because they were difficult to categorize. The size of this group is never more than 20% of the total consumers and is usually much smaller.

²Aluminum Foil (1962-66), Waxed Paper (1963-66), Liquid Detergent (1959-61), Facial Tissue (1958-61), and Headache Remedies (1959-61).

is that sample sizes may become small for certain cells when two or three sets of independent variables are simultaneously analyzed along with the dependent variable. This is a particularly vexing problem here because most demographic and household resource variables are inter-correlated. For example, high income households who rented and did not own a car were a very small percentage of all high income households. Initially, the data were analyzed individually for each explanatory variable. Then certain combinations of variables were considered together. In future studies, if larger samples are available, interrelationships between more sets of explanatory variables should be studied.

Household Resource Variables

The first two variables analyzed were home ownership and car ownership. Tables 1 and 2 give the results for each of the five product categories. The table entries are the percentage deal prone. For example, in the case of waxed paper, 29.5% of the households who owned a home were deal prone and 12.5% of the households who did not own a home were deal prone.

TABLE 1
HOME OWNERSHIP

Home Ownership	Product Category				
	Aluminum Foil	Waxed Paper	Headache Remedies	Liquid Detergent	Facial Tissue
Rent	30.9% ^a (81) ^b	12.5% (88)	22.7% (97)	29.6% (98)	23.4% (137)
Own	37.5% (120)	29.5% (152)	29.8% (151)	38.9% (229)	28.7% (230)

^aThe table entry is the percentage of aluminum foil buyers who rent a home and are deal prone.

^bNumbers in parentheses are total within-cell sample sizes on which each percentage is based.

TABLE 2
CAR OWNERSHIP

Car Ownership	Product Category				
	Aluminum Foil	Waxed Paper	Headache Remedies	Liquid Detergent	Facial Tissue
No Car	24.0% ^a (50) ^b	17.1% (70)	19.4% (67)	26.9% (67)	25.0% (88)
Car	38.4% (151)	25.3% (112)	30.0% (180)	38.8% (258)	27.3% (278)

^aSee Table 1, footnote a.

^bSee Table 1, footnote b.

Looking at the results in Tables 1-2, it appears that owning a car or a home makes a household much more deal prone. For every product category this result holds. These results are consistent with the predictions made in the previous section which indicated that home ownership and car ownership should lead to greater deal proneness.

One problem with analyzing the variables separately is that if one owns a car, one is also more likely to own a home. Thus, the observed effect may be due to one of the two variables and not the other. Table 3 studies the effect of the two variables jointly. The results show that except for facial tissue, owning both a home and a car results in the highest probability of being deal prone. The percentage deal prone is always higher when a household owned a car and a home than when it owned a car and rented. It is higher for 4 of the 5 products than when the household owned a home, but not a car. Thus, the effect is not due to only one of the two variables.

TABLE 3
CAR AND HOME OWNERSHIP

Home Ownership and Car Ownership	Product Category				
	Aluminum Foil	Waxed Paper	Headache Remedies	Liquid Detergent	Facial Tissue
No Car and Rents	21.2% ^a (33) ^b	14.0% (43)	22.9% (48)	28.6% (35)	18.3% (60)
No Car and Owns Home	29.4% (17)	22.2% (27)	10.5% (19)	25.0% (32)	39.3% (28)
Owns Car and Rents	37.5% (48)	11.1% (45)	22.4% (49)	30.2% (63)	27.3% (77)
Owns Car and Home	38.8% (103)	31.4% (105)	32.8% (131)	41.5% (195)	27.4% (201)

^aSee Table 1, footnote a.

^bSee Table 1, footnote b.

To get some idea of the magnitude of the effect that owning both a car and a home had on deal proneness, the following model was estimated:

$$D_{ij} = \alpha \beta_i \gamma_j \epsilon_{ij} \quad \begin{matrix} i=1, \dots, 4 \\ j=1, \dots, 5 \end{matrix} \quad (1)$$

where:

D_{ij} = percentage of deal-oriented consumers for product category j and consumer characteristic i .

α = average deal-orientation.

β_i = the effect of consumer characteristic i on deal-orientation.

γ_j = the effect of product category j on deal orientation.

ϵ_{ij} = the disturbance term.

Because product categories with very low deal proneness may have a lower absolute difference in deal proneness for a given household characteristic than product categories with high deal proneness, a multiplicative model was used. If one takes logarithms of both sides of equation (1), the effects (β_i and γ_j) can be measured using estimates calculated from standard analysis of variance formulas. The constraints are that

$$\sum_i \ln(\beta_i) = \sum_j \ln(\gamma_j) = 0. \quad (3)$$

The model is similar to the log-linear models described by Green et. al. (1977) and Bishop et. al. (1975). The exact estimates used are given in Hogg and Craig (1970, pp. 327-334).

Estimates of the model's β_i parameters are presented in Table 4. The results show that owning a car and a home had a 1.366 response, compared to not owning a car and renting, which had a .821 response. If we average across all the products so that the grand mean represents average deal responsiveness, we see that owning a car and a home increased deal responsiveness from 20.5% to 34.4%, a 67.9% increase. Owning either a car or a home, but not the other, increased deal responsiveness from 20.5% to 26.2%. It is clear, therefore, that owning both a car and a home greatly increases deal responsiveness, compared to not owning either or owning only a car or only a home.

³This constraint is the same as requiring $\prod_i \beta_i = \prod_j \gamma_j = 1$.

TABLE 4
CAR AND HOME OWNERSHIP

Grand Mean = 24.9%

Category	Response
No Car and Rents	.821
No Car and Owns Home	.932
Owns Car and Rents	.957
Owns Car and Home	1.366

Income

On theoretical grounds, household income level should be negatively correlated with deal proneness. Empirically, the opposite relationship may be observed because of the effects of confounding variables. Table 5 gives the results of income for three income levels -- low (\$0-5999) medium (\$6000-8999) and upper (\$9000 or more). The income categories are based on roughly 33% groupings. The results indicate that contrary to theoretical predictions, upper income households are more deal prone than low income households. For every product category except facial tissue, a higher percentage of upper income households are deal prone compared to low income households. The effect is not large in most categories, but it is persistent.

TABLE 5
INCOME

Income Level	Product Category				
	Aluminum Foil	Waxed Paper	Headache Remedies	Liquid Detergent	Facial Tissue
0-5,999	31.0% ^a (84) ^b	19.6% (97)	28.0% (100)	28.1% (114)	31.4% (137)
6,000-8,999	39.1% (64)	23.9% (71)	23.0% (87)	36.8% (114)	19.2% (125)
> 9,000	35.8% (53)	26.9% (52)	31.1% (61)	44.4% (99)	29.5% (105)

^aSee Table 1, footnote a.

^bSee Table 1, footnote b.

To isolate the effects of confounding variables, income was analyzed adjusting first for home ownership and then for car ownership. (Because cell sizes became very small, it was impossible to analyze income simultaneously with both home and car ownership). Tables 6-7 present the results. They indicate that if one adjusts for resources available, the income effect no longer persists (except for liquid detergent).

TABLE 6
HOME OWNERSHIP AND INCOME

Income Level and Home Ownership ^c	Product Category				
	Aluminum Foil	Waxed Paper	Headache Remedies	Liquid Detergent	Facial Tissue
0-5,999 and Owns Home	37.1% ^a (35) ^b	28.9% (58)	40.9% (44)	31.8% (66)	35.0% (62)
0-5,999 and Rents	26.5% (49)	13.6% (59)	17.9% (56)	22.9% (48)	29.3% (75)
6,000-8,999 and Owns Home	45.0% (40)	30.2% (53)	22.0% (59)	36.1% (85)	25.5% (85)
6,000-8,999 and Rents	29.2% (24)	5.6% (18)	25.0% (28)	35.7% (31)	10.0% (40)
9,000 or More and Owns Home	31.1% (45)	29.3% (41)	29.2% (48)	47.5% (80)	30.1% (83)

^aSee Table 1, footnote a.

^bSee Table 1, footnote b.

^cCertain categories were omitted because the sample sizes were too small.

TABLE 7
CAR OWNERSHIP AND INCOME

Income Level and Car Ownership ^c	Product Category				
	Aluminum Foil	Waxed Paper	Headache Remedies	Liquid Detergent	Facial Tissue
0-5,999 and No Car	21.6% ^a (37) ^b	17.6% (51)	20.0% (45)	21.1% (38)	25.9% (58)
0-5,999 and Owns Car	38.3% (47)	21.7% (36)	35.2% (54)	32.0% (75)	35.4% (79)
6,000-8,999 and Owns Car	37.5% (56)	23.3% (60)	25.0% (72)	37.5% (96)	19.0% (105)
9,000 or More and Owns Car	39.6% (48)	27.3% (44)	31.5% (54)	46.0% (87)	29.8% (94)

^aSee Table 1, footnote a.

^bSee Table 1, footnote b.

^cSee Table 6, footnote c.

To see this more clearly, Tables 8-9 present the β_1 parameter estimates for the model in equation (1). First, for income and home ownership, the highest deal-oriented category is low income households that own a home. The difference between rents and owns for all three income groups is dramatic. The income effect seems to disappear when rent or own is taken into account. Similarly for car ownership, the income effect is small compared to whether a car is owned.

TABLE 8
INCOME LEVELS AND HOME OWNERSHIP

Grand Mean = 26.3%

Category ^a	Response
0-5,999 and Owns Home	1.304
0-5,999 and Rents	.808
6,000-8,999 and Owns Home	1.151
6,000-8,999 and Rents	.661
9,000 or More and Owns Home	1.248

^aCertain cells have not been analyzed because of very small sample sizes.

TABLE 9
INCOME LEVELS AND CAR OWNERSHIP

Grand Mean = 28.2%

Category ^a	Response
0-5,999 and No Car	.748
0-5,999 and Owns Car	1.133
6,000-8,999 and Owns Car	.974
9,000 or More and Owns Car	1.213

^aSee the footnote to Table 8

Thus, higher income does not lead to deal proneness as would be concluded if the income effect were not adjusted for car and home ownership. If the effects of car and home ownership could be simultaneously partialled out, we would expect to find even stronger evidence that lower income households are more deal prone than higher income households.

CONCLUSIONS

This paper has used a specific model of buyer behavior to identify household variables which should affect deal proneness. The model assumed that households make inventory decisions the same way that firms do. The variables that affect the buying decision are holding costs, stockout costs, transaction costs, purchase price and usage rates. Household characteristics were linked to these cost variables, and predictions were made about which types of households should be deal prone. The predictions were then tested empirically.

The empirical results showed that household resource variables: car and home ownership, were strong predictors of deal proneness. 34.4 percent of the households that owned a car and home were deal prone. Only 20.5 percent of the households that did not own either a car or a home were deal prone. The effects of income were also analyzed and the results showed that upper income households were more deal prone. However, when income was adjusted for household resources, this effect became negligible.

REFERENCES

- Becker, Gary S. "A Theory of the Allocation of Time," Economic Journal, 75 (September 1965), 493-517.
- Bishop, Yvonne M.M., Stephen E. Fienberg and Paul W. Holland. Discrete Multivariate Analysis: Theory and Practice (Cambridge, Massachusetts: The M.I.T. Press, 1975).
- Blattberg, Robert C., Thomas Buesing, Peter Peacock and

Subrata K. Sen. "Identifying the Deal Prone Segment," Journal of Marketing Research (forthcoming, 1977).

Blattberg, Robert C. and Subrata K. Sen. "Market Segments and Stochastic Brand Choice Models," Journal of Marketing Research, 13 (February 1976), 34-45.

Frank, Ronald E., William F. Massy and Yoram Wind. Market Segmentation (Englewood-Cliffs, N.J.: Prentice-Hall, 1972).

Green, Paul E., Frank J. Carmone and David P. Wachspress. "On the Analysis of Qualitative Data in Marketing Research," Journal of Marketing Research, 14 (February 1977), 52-59.

Hillier, Frederick S. and Gerald J. Lieberman. Operations Research (San Francisco: Holden-Day, Inc., Second Edition, 1974).

Hogg, Robert V. and Allen T. Craig. Introduction to Mathematical Statistics (New York, NY: The Macmillan Company, Third Edition, 1970).

Kunreuther, Howard. "Why the Poor May Pay More for Food: Theoretical and Empirical Evidence," Journal of Business, 46 (July 1973), 368-83.

Michael, Robert T. and Gary S. Becker. "On the New Theory of Consumer Behavior," Swedish Journal of Economics, 75 (December 1973), 378-96.

Montgomery, David B. "Consumer Characteristics Associated with Dealing: An Empirical Example," Journal of Marketing Research, 8 (February 1971), 118-20.

Webster, Jr., Frederick E. "The 'Deal-Prone' Consumer," Journal of Marketing Research, 2 (May 1965), 186-89.

COMPARATIVE ADVERTISING: ISSUES AND PROBLEMS

Michael Etgar, State University of New York at Buffalo
 Stephen A. Goodwin, State University of New York at Buffalo

Abstract

The recent growth in use of comparative advertising has led to increased efforts to measure its effectiveness. Empirical efforts in this area should address seven important research issues. This paper identifies and expounds upon these issues and evaluates recent empirical studies of comparative advertising, particularly with respect to the way they have addressed these issues.

Introduction

The general research issue involving the communications-effectiveness of comparative advertising as compared to alternative message-presentation styles cannot be resolved until a large number of research questions have been tackled and answered. It is possible that the effectiveness of comparative advertising messages will depend substantially on the specific conditions under which the ad is communicated. In particular, interaction effects need to be identified and measured, and the nature of the advertisement, the message, the promoted product, the compared-to-product(s) and the respondents must be carefully and explicitly specified.

This paper deals specifically with issues that should be addressed when studying and researching the effectiveness of comparative communications vis-a-vis other types of message appeals. In order to accomplish this, we review seven studies which empirically addressed the effectiveness of comparative advertising, our focus being placed upon elicitation of major issues and the extent to which these studies come to grip with such.

All seven studies to be reviewed used an experimental design of one form or another. In these experiments, a group of subjects were exposed to comparative ads while other groups were exposed to some form of a non-comparative communication. The effectiveness of each type of message style was generally measured along a vector of effectiveness variables (typically linked to some part, or all, of a hierarchy-of-effects model) computed from respondents' paper and pencil responses.

In order to facilitate presentation, the issues and research findings from the above-mentioned seven studies are summarized below in Table 1. In this Table, the issues are listed horizontally while the seven studies are listed vertically in approximately chronological order. Cell entries reflect positions taken on an issue and/or whether or not the specific issue was addressed.

TABLE 1

	Ogilvy and Mather (1975)	Golden (1976)	Wilson (1976)	Prasad (1976)	Mazis (1) (1976)	Mazis (2) (1976)	McDougall (1976)
<u>I. Product/Brand Characteristics</u>							
1. Product class perception of utility	Not considered	Not considered	Not considered	Not considered	Not considered	Not considered	Not considered
2. Product type used	Health & beauty aids	Deodorants	Automobiles, beer, cat food, credit cards, deodorant, mouthwash, toothpaste, soap	Cameras	Supermarkets, fast food restaurants	Supermarkets, fast food restaurants	Deodorants, laundry detergents, bleach
3. Product class salience	Not considered	Not considered	Not considered	Not considered	Not considered	Not considered	Not considered
4. Advertised products' life cycle stage	Maturity	Introduction/maturity	Maturity	Introduction	Maturity	Maturity	Introduction/Maturity
5. Sponsor's market position	Not considered	Considered ¹	Not considered	Not considered	Considered ¹	Considered ¹	Not considered
<u>II. Characteristics of Compared-to-Brands</u>							
1. Named or not	C vs NC vs Control ²	C vs NC	C vs BX	C vs BX	C vs BX, C vs Control, NC vs Control	C vs NC	C vs BX

¹ Manipulated

² C = Comparative ad, NC = Non-comparative (supportive) ad, BX = "Brand X" ad

TABLE 1 (Continued)

	Ogilvy and Mather (1975)	Golden (1976)	Wilson (1976)	Prasad (1976)	Mazis (1) (1976)	Mazis (2) (1976)	McDougall (1976)
2. Market position	Unknown	Second leading brand	The leading brand	The leading brand	Not considered	Not considered	Two leading brands ³
3. Similarity	Not considered	Not considered	Not considered	Not considered	Not considered	Not considered	Not considered
4. Number of brands compared to	Unknown	One	One	One	One	One	Two ³
<u>III. Product/Brand Attributes</u>							
1. Number of attributes compared	Unknown	Two ³	One	Four ³	One	One	Four ³
2. Salience of attributes used for comparison	Not controlled	Checked after selection	Not controlled	Not controlled	Not controlled	Not controlled	Not controlled
<u>IV. Respondents</u>							
1. Type	450 female heads of households	594 college students	80 college students	202 college students	162 college students	87 college students	225 female heads of households
2. Degree of knowledge of product class/brands	Not controlled	Not controlled	Controlled for	Controlled for	Not controlled	Not controlled	Not controlled
3. Usage rates	Not controlled	Controlled for	Not controlled	Not controlled	Not controlled	Not controlled	Not controlled
4. Brand loyalty	Not controlled	Controlled for	Not controlled	Not controlled	Not controlled	Not controlled	Not controlled
5. Personality traits	Not controlled	Not controlled	Not controlled	Not controlled	Not controlled	Not controlled	Not controlled
<u>V. Ad Design</u>							
1. Media vehicle	Broadcast only	Print only	Print only	Print only	Print and audio ⁴	Audio only	Print only
2. Number of ads presented	Eight (portfolio) ³	One	Seven (portfolio) ³	Several (portfolio) ³	Fifteen (portfolio) ³	Eight (portfolio) ³	One
3. Frequency	One presentation	One presentation	One presentation	One presentation (two measurements)	One presentation	One presentation	One presentation
<u>VI. Message Characteristics</u>							
1. Substantiation	Unknown	Manipulated	Unsubstantiated	Unsubstantiated	Unsubstantiated	Unsubstantiated	Manipulated
2. One-sided or two-sided	One-sided only	One-sided only	One-sided only	One-sided only	One-sided only	Two-sided ⁴	One-sided only
3. Position of the claim in the ad (recency-primacy)	Not considered	Not considered	Not considered	Not considered	Not considered	Not considered	Not considered

³Not manipulated⁴Manipulated

TABLE 1 (Continued)

	Ogilvy and Mather (1975)	Golden (1976)	Wilson (1976)	Prasad (1976)	Mazis (1) (1976)	Mazis (2) (1976)	McDougall (1976)
VII. Testing & Methodology							
1. Independent variables	Type of message appeal (C - NC)	C - NC ads, Advertising competitive position, Claim substantiation, Copy theme	Type of message appeal (unsubstantiated, subjective C)	BX vs C	Media (audio & print), Type of message appeal (BX or C), Type of sponsor (supermarket or fast food chain), Market position of sponsor (leader or follower)	Type of message appeal (C or NC), Message-sidedness (1 or 2), Market position of sponsor (leader or follower)	Type of message (C vs BX), Substantiation (substantiated vs unsubstantiated)
2. Dependent variables	Believability, Confusion perception of important differences in brands advertised, Unaided brand identification, Persuasiveness of ad (pre- & post-exposure brand choice differences)	Purchase intentions, Relative importance of product attributes mentioned in ad	Ad informativeness, Believability, Offensiveness, Interestness; Product quality, Trustworthiness of sponsor	Unaided brand recall, Perceived credibility of claim, Perception of sponsor's competitive position, Message recall one week after exposure (& claim recall but aided)	Unaided recall, Aided recall, Over-all brand attitude measures, Dynamism, authoritativeness & character of message	Attitude & belief scale, Unaided & aided recall, Free response protocols, Dynamism & character of message	Measures of reliability & helpfulness
3. Pre-tests/manipulation checks	None	None	None	None	None	None	None
4. Covariates/block variables	Considered only product class users for analysis	Brand loyalty (toward advertised & compared-to-brand)	Considered only product classes which were familiar & used by subject population	Prior preference of compared-to-brand established, Kelman & Cohler's test for cognitive clarity	None	None	Users and non-users
5. Analytical technique	Difficult to determine	Analysis of covariance, Analysis of variance	Analysis of variance (split plot design), Scheffe multiple comparison test	Goodman's contingency table analysis, Mann-Whitney U test, Chi-square analysis	Analysis of variance, Multiple comparison tests	Analysis of variance, T-tests	Regression

Discussion of the matrix is row-wise, reflecting our primary interest in the issues involved. In each instance, we discuss the research topic involved and then the approaches taken (if any) by specific authors. A critical review of these approaches is presented and in some cases suggestions for improvement are offered.

The issues explored pertain to five major dimensions: the characteristics of the brands which appear in the ad, the audience toward whom the ad is directed, the message design utilized, other factors relating to the ad (media utilized, multiple exposures, etc.), and the methodology used for testing and evaluating ad effectiveness.

Product Class/Brand Characteristics

It is generally accepted among marketing scholars that (regardless of the type of message-style utilized) an advertisement should develop and/or reinforce audience cognitions, attitudes, and intentions which will ultimately lead to trial and (hopefully) repeat purchase of the advertised brand. Whatever the reference point, it would seem obvious that certain characteristics of the brands presented in a communication can have crucial implications for message design. In evaluating the effectiveness of one type of message design versus another, such characteristics should be analyzed.

While in non-comparative (typically referred to as supportive) ads, designers have to be concerned only with the characteristics of the promoted brand, in comparative communications they have to consider also those of the compared-to-brands as well as the dimensions of comparison (e.g., product attributes) to be used in the ad.

Characteristics of the Advertised Brand

Product Typology. A major rationale advanced in favor of comparative advertising is that naming names of specific brands to which the advertised brand is compared, provides the consumer with more factual information and consequently aids him in making more rational brand choices (c.f., Wilkie and Farris, 1975). Consequently, comparative advertising is primarily applicable to messages which compare specific measurable product attributes such as weight, price, speed, etc., and therefore may be especially applicable for products which are primarily bought because of their functional attributes; namely, their performance along measurable dimensions, or their capacity to perform specific measurable tasks.

A large number of products, however, are bought not because of their task-related capacities, but because of the psychological or social benefits which they promise. The ability of an ad to compare and contrast performance of different brands in providing elusive benefits is more restricted. While one can establish or claim with a reasonable degree of credibility (of course, the Substantiation Doctrine of 1972 would apply in this instance; FTC Report, 1972) that one automobile brand, for example, has a better gas mileage performance than another brand, it is difficult to extend the argument and claim that one auto confers more status to its user than another brand.

The effectiveness of comparative advertising may therefore depend on whether the advertised product provides primarily functional, utilitarian benefits or social/psychological ones. However, none of the seven studies summarized in Table 1 have explicitly considered this issue by integrating it into the experimental designs, or comparing comparative advertising effectiveness for different product classes (differing along a functional-social continuum).

Product Life Cycle Stage. While some researchers used existing brands in their experiments, others relied on new (typically fictitious) brands as the advertised stimulus. The use of existing and well-known brands in an experiment raises the issue of subjects' prior experience with and current attitudes toward such brands. It is important to recognize that such predispositional information should be controlled for. In many cases, consumers have well established perceptions and opinions about specific brands which may have required prolonged promotional effort involving the encoding of numerous messages over substantial periods of time. Not controlling for such "contamination" would tend to complicate inferences drawn from an experimental study.

Additionally, it is reasonable to expect that "one exposure" experiments (which was the case for all seven published studies) could create insignificant marginal changes in respondents' attitudes and perceptions of well-known brands (in at least a "maturity" stage of the product life cycle), regardless of the message design utilized.

Thus, unless subject predispositional variables are taken into account (e.g., through analysis of covariance or use of a treatment-by-levels type design; Lindquist, 1956), it is preferable to develop messages for new fictitious brands.

Characteristics of The Compared-to-Brands

Naming of Compared-to-Brands. As defined by Wilkie and

Farris (1975) a comparative advertisement must either specifically name the compared-to-brand(s) or refer to them in such a way as to leave no doubt regarding what they are. Additionally, the comparison must involve specific attributes of the brands.

If we utilize Wilkie and Farris' definition as a reference point, then any indirect comparison should be general so as to preclude consumer recognition of a specific comparison (e.g., Brand A lasts longer than other leading brands, as opposed to Brand A lasts longer than the leading seller).

The so-called "supportive" or non-comparative ad may also be viewed, at the most general level, as a comparative ad. Even where there is absolutely no allusion to a competitive offering, where the message communicates only information about the promoted brand, is it not fair to say that the consumer is likely to mentally make a comparison with available and known alternatives? Thus, we have a continuum of message approaches which include truly comparative communications, Brand X ads, and supportive ads. In operationalizing communication styles, care must be taken to ensure that the experimental ads truly reflect the underlying conceptual structure. For example, in the Prasad (1976) study, it is extremely likely that a number of subjects knew which brand was the "leading brand" of movie cameras. If this assumption is tenable, then Prasad has not, in fact, manipulated comparative versus Brand X ads. In effect both conditions presented comparative communications (particularly for knowledgeable respondents).

In reviewing the seven studies, we see that there is some doubt as to the existence of isomorphic relationships between concept and variable; and this is particularly the case for the so-called Brand X advertisement.

Market Position of Compared-to-Brand(s). The market position(s) of the compared-to-brand(s) are important to consider because such position(s) may affect the consumer's frame of reference as it relates to the evaluation of the brand promoted. One strategy commonly found is to compare the promoted brand to the leading brand in the pertinent generic product class though its superiority is far from established. One assumed advantage of using a leading brand for comparison is that it can potentially reduce misidentification of the promoted brand and misunderstanding of the corresponding message as well as attract the attention of current users of the leading brand and thus, at a minimum, lead to increased likelihood of greater awareness for the advertised brand.

Other researchers disagree. Kershaw and Tannenbaum (1976) argue that where both the promoted and the compared-to-brands are relatively unknown to the audience, the viewers may confuse the brands, possibly leading to the paradoxical situation whereby a comparative ad improves consumer attitude toward (and ultimately sales of) the compared-to-brand.

Still other opinions reflect the belief that use of market leaders for comparison may be a futile exercise in increasing awareness. While a greater number of consumers might be attracted to the ad, these people may be much less prone to change their attitudes and purchase behavior patterns due to their brand loyalty; further, such brand loyal consumers may be more likely to disbelieve the message claims made by the "new" (or small, in terms of market share) brand. Perhaps, then, anchor brands should be other than the market leader.

So far, none of these arguments has been validated by actual research. Even though Golden (1976) and Mazis (1976) have manipulated the market position of the sponsored brand, to date, none of the published studies have manipulated the market position of the compared-to-brands. Clearly this area may be very important

vis-a-vis practical marketing strategy implementation.

Similarity of Brands. In many generic product classes, subgroupings exist. For example, even though economy, luxury and sports cars belong to the same general product class of automobiles, consumers may use different evaluative criteria and hold different evoked sets in each group.

Comparison claims between brands from different subgroups may be tenuous. There are two points to be made in this context; first, consumer perceptual/attitudinal evaluation of ads which make "dissimilar comparisons" may be more negative than for comparisons involving brands likely to be included in the typical evoked set; second, the legal ramifications of such a strategy are not clear. While the claims made may be literally true (e.g., VW has more trunk space than Cadillac Coupe de Ville, better gas mileage than Mercedes 220 SL, etc.) the impression made may be false. Thus, the advertiser considering a comparative message appeal between dissimilar items runs the risk of FTC involvement on the basis of advertising deception. None of the published studies have dealt with this issue. It merits empirical inquiry.

Number of Brands Compared. Some comparative ads include only one compared-to-brand; others use several. On an a priori basis, there is little behavioral evidence to indicate which approach should be more effective, if any. However, it may be reasonable to expect that a substantial increase in the number of brands compared can add to communication noise (e.g., information overload), while a modest comparison set may improve the credibility of the message. The bulk of the reported studies utilize only one brand for comparison. The exception is McDougall (1976) who used two. Experimental manipulation of this variable has not, in any case, been operationalized.

Product Attributes on Which Brands Are Compared

Number of Attributes Presented

Comparative ads found currently in the media reflect a diversity of approaches. While some restrict discussion to one product characteristic, many make comparisons on multiple dimensions, believing that more information lends additional credibility concerning the claim (or in order to be consistent perhaps with a particular positioning strategy which management wishes to invoke). The issue of course is to find out, other things being equal, the differential effects (if any) due to presentation of differing amounts of attribute information.

Perhaps the "information overload" (Jacoby, et al., 1974) framework which suggests that consumer evaluation of ad and brand advertised conforms to some kind of inverted u-shaped function of amount of information provided in the communication is useful here.

If this argument is tenable, a comparative ad which uses one or a small number of attributes as comparison points may have more "positive" effects than similar appeals making a greater number of attribute comparisons; or, perhaps an intermediate amount of brand characteristic information would be "best". Most of the studies in Table 1 have used one or two product attributes for comparison: two (Prasad, 1976; and McDougall, 1976) have used four characteristics. None of these studies, however, have operationalized an "amount of attribute information" variable and attempted to compare the impact of change in this variable on communicative effectiveness.

Saliency of Brand Attributes

The effectiveness of a comparison between two or more brands on a specific attribute dimension in terms of

consumer evaluative response may depend very much on the relative importance (perception of saliency) of the attributes for the targeted audience. If the discussed characteristics provide important benefits to the consumer, then the comparison may be interesting and informative, and likely to make a positive impact. Consumers who are only marginally interested in the brand attribute(s) discussed may not be influenced at all (or even negatively influenced).

Consequently, pretesting the saliency of brand attributes should be a basic initial step in any research design for testing effectiveness of comparative communications. Yet none of the published studies reported in Table 1 have selected product attributes according to such a pretest; therefore, it may be that, for the subject population studied, marginal or non-salient attributes were used for comparison. Golden (1976) did measure product attribute importance in her study, but apparently only after selection for the experiment. She found that selected attributes were rated as important; but she had no way of knowing if others, not included, were relatively more important.

Respondents

Type of Respondents

Researchers studying effectiveness of comparative advertising have employed two distinct groups of subjects: students and female heads of households (housewives). Use of adult housewives has generally been lauded due to the ostensible increase in external validity of the study results. Furthermore, it may be likely that housewives are more concerned with product classes which have been utilized in the various studies, and consequently are more involved in the task presented them regarding assessment of ad and brand advertised. Still, use of students as subjects may be entirely acceptable and consistent with improved external validity if product classes advertised are salient to them (i.e., they purchase and consume brands from such product classes). Two of the studies in Table 1 (Ogilvy and Mether, 1975; and McDougall, 1976) have used female heads of households as subjects; the others have used students. Only Wilson (1976) and Prasad (1976) explicitly measured product/brand relevance explicitly.

Knowledge of the Product Class/Brands

In order to derive evaluative conclusions from a message presented by a comparative communication (or any type of appeal, for that matter), the consumer will draw upon knowledge gained about the compared-to-brands. Certainly, lack of such knowledge will have some impact on the evaluation process; at the extreme, evaluations of the comparison claims may be meaningless to a particular viewer heretofore unknowledgeable about the product class and/or brands promoted and compared. Of the studies reviewed, only Wilson's checked for degree of prior consumer knowledge.

Usage Rates. A useful surrogate for product knowledge, perhaps, is usage rate which details the extent to which the consumer has been buying or using various brands from the pertinent product class(es). Usage rates were not incorporated in any of the studies in Table 1.

Brand Loyalty. Related to usage is the notion of brand loyalty. Consumers who exhibit substantial brand loyalty to compared-to-brand(s) in comparative communications may be likely to resent the comparative claim(s), perhaps in response to cognitive dissonance. Whether or not, in principle, this cognitive response is general, the point remains that respondents should be measured on this basis to permit an examination of such effects. As with the above suggested measures, brand loyalty measurements

permit increased control of error variation, and can be utilized in experimental designs via covariance analysis or directly in the form of a leveling variable.

Brand loyalty was incorporated in two studies: Golden (1976) and McDougall (1976). Golden measured brand loyalty by subjects' reported relative frequency of purchase of the advertised and compared brands, and used it as a covariate in an analysis of variance design. McDougall measured brand loyalty by assessing subjects' degree of agreement with a semantic differential statement with endpoints "loyal to the (specific) brand" and "different brands are purchased", or "don't buy". He then compared the means of responses regarding subjects' opinions as to the reliability and helpfulness of the ads between loyal and non-loyal subjects, by using a simple t-test.

It should be mentioned, too, that Prasad (1976) measured his subjects regarding the number who already (prior to the test) preferred the compared-to-brand (in this instance, Kodak XL movie camera); he found, via nonparametric tests, that there was no evidence of "selective recall" between Kodak-preferred and non-Kodak-preferred groups.

Personality Traits

While we do not advocate use of general personality inventories (Kassarjian, 1971), we do feel that certain carefully selected measures may shed additional light on consumer response to comparative vis-a-vis non-comparative communications. In particular, those traits which pertain to the ability and motivation of the consumer's processing of information varying in terms of uncertainty, ambiguity, etc., are of some interest. While none of the six authors incorporated personality trait measures in the external designs a couple can be mentioned here. The Mehrabian-Russell Arousal-Seeking Tendency Scale (Mehrabian and Russell, 1974) which measures the general tendency or proneness of curiosity-seeking/variety-seeking behavior is one example. Consistent with a point mentioned by Wilkie and Farris (1975) that comparative ads are likely to be perceived as more novel than "ordinary" supportive message appeals, we should find that consumers high on arousal-seeking tendency would evaluate comparative ads more positively than those consumers "low" on arousal-seeking tendency. A second example is Kelman and Cohler's (1967) test for "cognitive style". This test identifies two major styles regarding consumer's characteristic mode of resolving cognitive uncertainty: clarifiers (people who react "positively" to uncertainty by seeking clarification, additional information, etc.) and simplifiers (people who simplify their environments by keeping out obtrusive cognitions, and tend to avoid uncertainty at "all costs"). Assuming that the information encoded in comparative communications may arouse the cognition of uncertainty or ambiguity (at least as compared to a more traditional supportive communication), it may be reasonable to suppose that clarifiers will explore comparative communications more thoroughly, and generally evaluate them more favorably (perhaps be affected more by them), than would be the case for simplifiers.

Other Design Considerations

Medium

The studies presented in Table 1 cover a good range of media-vehicles. Several experiments invoked mock print advertisements, others used radio and/or television (video-taped) ads. McDougall (1976) suggests that for comparative advertising, print media may have substantial advantages over audio-visual media because comparative advertising includes relatively more information (data not only about the promoted brand but also about additional brands) and correspondingly more complex messages, and thus may be communicated more effectively via

the print medium where the consumer can spend more time with the message, and do so at his/her own pace. Audio-visual approaches automatically restrict message-consumer interaction time.

Two of the summarized studies (Mazis, 1976; and Ogilvy and Mather, 1975) have operationalized their communications via audio-visual or only audio vehicles. The remaining studies used print medium only. Mazis also compared print versus radio media. In this study he found insignificant differences on major response variables between message modalities, shedding some light on the previously mentioned opinion that comparative advertising is likely to be more "effective" in print rather than audio-visual media. The issue is still open, however, for future study.

Number of Ads Presented

Some of the researchers exposed their subjects to only one advertisement and compared the results between conditions using essentially factorial experimental designs. Others exposed subjects to several ads, all of which were studied (Wilson, 1976) or only the one which was tested (e.g., Prasad, 1976; here the rationale of using a portfolio approach was to create a more "natural" format and simulate a "real life" exposure situation).

It is possible that the portfolio approach inadvertently creates unnecessary demand characteristics; subjects' attention is spread out, a memory set is likely to be aroused, etc. It is important to come to grips with the potential trade-off between "real-life" simulation and confounding demand characteristics. (We recognize that if the primary response domain pertains to memory, a portfolio approach is virtually mandatory: more on this in the last major section of the paper.)

Frequency/Intensity of Exposure

Cognitive response to advertising may be affected by the frequency of message exposure. Repetition effects (Sawyer, 1973; Ray and Sawyer, 1971) deserve careful study, but to date the experiments published have invariably been single-exposure. It is difficult to assess those studies that have focused upon recall (aided or unaided) measures in evaluating comparative versus non-comparative ads when only one single exposure has taken place. Replication designs ought to be considered by future researchers; furthermore, recall/memory measures should be taken at realistically spaced time intervals rather than immediately (or shortly) after advertisement exposure. All of the studies measuring recall suffer from this potential problem, although Prasad (1976) readministered his measures to subjects one week after his single-exposure.

Other Message Characteristics

Message characteristics which may affect cognitive, affective, or behavioral responses to communications include the issues of substantiation (factual or subjective claims), sidedness of the message, and the relative position of the comparative claims within the advertisement.

Substantiation

Several critics of comparative advertising (Kershaw, 1976; Chevins, 1975) have suggested that when the comparisons include generalized, non-substantiated statements, the consumer may become misinformed as well as confused. Claims of superiority, for example, which are not supported by independent (or company) tests may reduce the credibility of the message. McDougall (1976) tested for the relative effectiveness of use or non-use of substantiated claims across various message appeals. Using separate regressions for comparative substantiated, noncomparative

substantiated, comparative unsubstantiated, and non-comparative unsubstantiated messages, with measures of ad reliability and helpfulness as dependent variables, McDougall found that, overall, substantiated claims receive greater acceptance in terms of reliability and helpfulness; non-comparative substantiated claims were found to be superior to all of the others. Golden (1976) also manipulated claim substantiation. She found no differences in message-appeal effect.

Wilson's study (1976) also bears on this issue in that all of his ad manipulations involved subjective (unsubstantiated) claims. He found that unsubstantiated comparative ads were significantly perceived as more confusing and less believable than their unsubstantiated non-comparative counterparts.

One-Sided versus Two-Sided Appeals

In his working paper, Mazis (1976) provides an interesting perspective regarding sidedness of message appeals. Drawing upon work done by Hovland, et al., (1949), and integrating McGuire's Inoculation Theory (1961), Mazis presents a conceptual framework for assessing message sidedness with respect to comparative or non-comparative advertising. Specifically, Mazis points out that comparative communications can be constructed along conventional lines by mentioning only positive comparisons in the advertised products favor (one-sided comparative ad), or alternatively by mentioning both positive and negative facts about the promoted brand (two-sided comparative ad).

From the managerial perspective, it may be the case that such communications are not only more believable, but that the overall image of the brand (and perhaps the company sponsoring it) is enhanced due to the "refreshing" glimpse at honesty. It is plausible to assume that most consumers would applaud an attempt at "telling it like it really is" rather than only mentioning positive benefits accruing to brand consumption. From the public policy point of view, such two-sided comparative ads may provide even more objective information upon which the consumer may make a brand choice. This is entirely consistent with the stance taken by the FTC.

Operationally, of course, there is always the chance that such a communication problem may win an ad agency an Effie, but lose the company precious sales and market share. It deserves careful study. Mazis initiates such a study in his second experiment reported in the above-mentioned working paper. The results are clouded, unfortunately, by an unrealistic operationalization of the negative appeal in the two-sided communication; nonetheless, he has made an important first step in studying the interesting, and potentially useful (managerially and public policy-wise) message-appeal strategy.

Relative Positioning of the Claim(s)

In rounding out this section, we mention that effectiveness of comparative advertising communications may depend, in part, on the positioning of the comparative claims (e.g., at the beginning, the middle or end of the ad). None of the seven studies explored this issue, although admittedly it is a minor point in contrast to some of those previously discussed. Also related to this "recency-primacy" issue is the question of whether the compared-to-brand(s) might be mentioned first or last in the message.

General Issues Pertaining to Testing and Methodology

Independent Variables

Manipulations of message appeals require close attention to proper operationalizations. Accordingly, scrutiny of the isomorphism between concept and variable is

imperative. Where correspondence is considerably less than 1 to 1, the researcher is not in a position to make conclusive inferences with respect to the response effects. Thus, it is important that the researcher fully understands what is meant by "comparative advertisement", "Brand X" ad, and "supportive" communication so that reasonable working definitions can be operationalized. The same is true whatever independent variable is selected for study.

We have seen, for example, that there is some doubt as to whether or not Prasad's (1976) "Brand X" ad was, in fact, truly perceived as a Brand X ad or rather a standard comparative ad (at least for some of the subjects in his study). As a general rule, it is advisable to thoroughly pretest these manipulations, utilizing different subjects from the same population. Such pretesting is the exception rather than the rule in published studies.

Dependent Variables

Without getting into the debate as to what constitutes reasonable measures of advertising effectiveness, we can say that perhaps the majority of advertising academics and certainly practitioners are comfortable with some sort of hierarchy of effects model (e.g., Lavidge and Steiner, 1961).

In such models the consumer is viewed as responding to advertising stimuli according to a hierarchy such as "awareness", "comprehension", "interest", "liking", "intention", and "behavior". As pointed out by Lavidge and Steiner (1961), such a portrayal of response to stimuli covers three stages: cognitive - affective - and conative (behavioral). The advertising researcher, over the years, has developed a variety of measures which more or less tap dimensions of each of these three stages. Depending upon the purpose of the ad, whether copy pretesting or market post-testing is the primary focus, combinations of these measures are gathered.

The seven studies summarized have utilized all or part of such hierarchical responses. Some (Prasad, 1976; Mazis, 1976; Ogilvy and Mather, 1975) stressed memory or recall ability occasioned by the various manipulated advertisements. Such measures reflect the researcher's concern for assessing the cognitive dimension. Others utilized a battery of questions designed to shed light on the entire spectrum, with the exception of memory measures (Wilson, 1976).

However, both evaluations of the ad itself and the brand advertised should be measured. Some of the studies looked at one or the other (e.g., Golden, 1976; McDougall, 1976), but not both.

Pretests and Manipulation Checks

It is extremely important that the researcher verify that his/her manipulations are in fact being perceived in a manner which is consistent with his/her objectives. This is true of independent variable operationalizations as well as dependent measures (scales) administered and gathered.

Thus, pre-tests/manipulation checks are crucial to ensure internal validity. While such checks can be made after the experiment is completed by using the same subjects, it may be more advisable to perform them before the experiment by using different subjects drawn randomly from the same population from which the experimental respondents will be selected. Such a pre-test strategy mitigates against likely demand characteristics operating (Sawyer, 1975). Thus, use of scale which is not discriminating will yield statistically insignificant results when, in fact, such results may exist. Golden's study suffers in this respect. All her subjects regardless of experimental condition appeared to generally use

the same small range on her intention scale. Perhaps the scale was not explained adequately; perhaps an 11-point subjective probability scale should have been utilized instead of the 7-point scale which was used; perhaps the independent variables were operationalized poorly.

Another related point concerns the nature of the measurement level of the scales used to tap dependent responses. Unless such scales are thoroughly examined and pre-tested so as to permit an assumption that they are interval-scaled, it is difficult to argue in favor of parametric analytical techniques. And, since parametric techniques (e.g., analysis of variance, regression) permit assessment of interaction effects (whereas assessment of such by use of non-parametric techniques is not clear cut), it behooves the researcher to strive for development of dependent measures which can be assumed to be at least interval scaled. This requires the discovery that subjects perceive distances between each number in a given scale as reflective of approximate equal psychological distances.

Covariates

When a researcher does not collect predispositional measures which presumably relate to informational processing of advertising stimuli, error variance may be unnecessarily large, resulting in non-significant differences.

To reduce error variation (or adjust for extraneous sources of error), the researcher might design experiments which include important predispositional variables (e.g., treatment-by-levels designs (also referred to as randomized block designs); see Lindquist, 1956; or Winer, 1971) as one (or more) of the factors. Alternatively, he might consider the use of such predispositional variables as covariates, where any contaminating influences may be statistically adjusted for. The Golden study is illustrative of this approach: she gathered a measure of brand loyalty which she used to statistically adjust responses to treatment combinations. Clearly, assessment of main and interaction effects (as well as the various multiple comparisons which may be of interest) can be clouded by failure to reduce reasonable sources of confounding error variation (e.g., product class usage experience, brand loyalties, evoked sets, related personality traits).

Summary and Conclusions

The purpose of this paper has been to explore the various issues involved in assessing the communications effectiveness of comparative advertising. Using seven recent studies as reference points, we have endeavored to point out the complexities of comparative advertising research. No matter what the scope of an individual study, there are several crucial considerations necessary. These involve the advertised product class (brand) characteristics (e.g., product class, salience, life cycle stage), the characteristics of the compared-to-brand(s) (e.g., market position(s), number of brands compared), product class (brand) attributes on which comparisons are made (e.g., number of attributes and their salience), the subjects (e.g., type, usage rates, brand loyalty, personality traits), the advertisement design (e.g., media-vehicle utilized, frequency/intensity of ad exposure), other message characteristics (e.g., use of substantiation, sidedness of appeals), and methodology used (e.g., experimental designs, manipulation checks, covariates).

The seven studies addressed only some of the research questions identified in this paper. Notably, there is an important gap in that the bulk of the studies failed to experimentally manipulate (or control error variation via covariance analysis) major product-, subject-, and situational-related variables. While it is true that exploration of all research questions raised in this

paper are beyond the scope of any individual study, we suggest that future research should address the specific issues presented.

Careful attention to the identified issue areas should contribute to the development of body of knowledge which could then provide answers as to when, if, and how comparative advertising can be used.

References

- A. G. Chevins, "A Case for Comparative Advertising," Journal of Advertising, 4:2(1975), 31-36.
- Federal Trade Commission Report. "Staff Report to The Federal Trade Commission on The Ad Substantiation Program Together with Supplementary Analysis of The Submissions and Advertisers' Comments," Committee on Commerce, U.S. Senate, Washington, D.C., July, 1972.
- L. L. Golden, "Consumer Reactions to Comparative Advertising," in B. B. Anderson, Advances in Consumer Research, Volume III (Chicago: Association for Consumer Research, 1976), 63-67.
- C. I. Hovland, A. A. Lumsdaine and F. D. Sheffield, Experiments on Mass Communication (Princeton: Princeton University Press, 1949).
- J. Jacoby, D. E. Speller and C. A. Kohn, "Brand Choice Behavior As A Function of Information Load," Journal of Marketing Research, 11(1974), 63-69.
- H. H. Kassarjian, "Personality and Consumer Behavior: A Review," Journal of Marketing Research, 8(November, 1971), 409-418.
- H. Kelman and R. Cohler, "Test for Cognitive Style," (Reported and adapted) in D. Cox, Risk-Taking and Information Handling in Consumer Behavior (Cambridge: Harvard University Press, 1967).
- A. G. Kershaw and S. I. Tannenbaum, "For and Against Comparative Advertising," Advertising Age, (July 5, 1976), 25-28.
- R. J. Lavidge and G. A. Steiner, "A Model for Predictive Measurements of Advertising Effectiveness," Journal of Marketing, (October, 1961), 59-62.
- E. F. Lindquist, Design and Analysis of Experiments (Boston: Houghton Mifflin, 1956).
- M. B. Mazis, "A Theoretical and Empirical Examination of Comparative Advertising," Unpublished paper, University of Florida, 1976.
- G. H. G. McDougall, "Comparative Advertising: An Empirical Investigation of its Role in Consumer Information," Working Paper #167, School of Business Administration, The University of Western Ontario, London, Canada, December, 1976.
- W. McGuire and D. Papageorgis, "The Relative Efficacy of Various Types of Prior Belief-Defenses in Producing Immunity Against Persuasion," Journal of Abnormal and Social Psychology, 62(May, 1961), 327-337.
- A. Mehrabian and J. A. Russell, An Approach to Environmental Psychology (Cambridge: MIT Press, 1974).
- Oglivy and Mather. "The Effects of Comparative Television Advertising That Names Competing Brands," Research Report, 2 East 48th Street, New York, 1975.

V. K. Prasad, "Communications-Effectiveness of Comparative Advertising: A Laboratory Analysis," Journal of Marketing Research, 13(May, 1976), 128-137.

M. L. Ray and A. G. Sawyer, "A Laboratory Technique for Estimating the Repetition Function for Advertising Media Models," Journal of Marketing Research, 8(February, 1971), 20-29.

A. G. Sawyer, "The Effects of Repetition of Refutational and Supportive Advertising Appeals," Journal of Marketing Research, 10(February, 1973), 23-33.

A. G. Sawyer, "Demand Artifacts in Laboratory Experiments in Consumer Research," Journal of Consumer Research, 1(March, 1975), 20-30.

W. L. Wilkie and P. W. Farris, "Comparison Advertising: Problems and Potential," Journal of Marketing, 39(October, 1975), 7-15.

R. D. Wilson, "An Empirical Evaluation of Comparative Advertising Messages: Subjects' Responses on Perceptual Dimensions," in B. B. Anderson, Advances in Consumer Research, Volume III (Chicago: Association for Consumer Research, 1976).

B. J. Winer, Statistical Principle in Experimental Design, 2nd edition (New York: McGraw-Hill, 1971).

MEDIATING EFFECTS OF COGNITIVE RESPONSES TO ADVERTISING ON COGNITIVE STRUCTURE¹

Jerry C. Olson, The Pennsylvania State University
Daniel R. Toy, University of Washington
Philip A. Dover, Dartmouth College²

Abstract

This paper reports results of an initial study of the mediating effects of cognitive responses to advertisements on selected elements of cognitive structure. To create variation in cognitive response, consumers were shown one of several new product ads that varied in terms of product price and the presence or absence of information regarding additional (non-price) attributes of the advertised product. Type of cognitive response was sensitive to both aspects of ad content. As expected, counterargument and support argument responses were both strongly related to several cognitive structure elements.

Introduction

The measurement and explanation of persuasive communications impact has occupied the attention of a wide variety of academic, industry, and public policy researchers. In literally hundreds of studies attempts have been made to model the effects of the persuasive communication process and its subsequent influence on consumer behavior. The importance of this issue for marketing and public policy is evidenced by the more than \$28 billion spent in the United States on advertising in 1976.

The lack of a clear understanding of the communication process may be attributed to several factors, the broadest and most critical of which may be the absence of a vigorous, process-oriented, empirical research tradition. The dominant research paradigm for communication research involves the measurement of some dependent variable (attitudes, opinions, sales) following exposure to a persuasive communication. The independent variables typically manipulated in these studies are usually concerned with characteristics of the source or channel of the message, the message itself, or the receiver of the message (Mitchell & Olson, 1977). The purpose of such studies is to identify generalizable relationships between the presumed goal of the persuasive communication (e.g., attitude change) and the communication variable of interest. For the most part, these simplistic attempts have not been successful (cf. Fishbein & Ajzen, 1975).

In contrast to this generally atheoretical approach to studying communication effects are two streams of research based upon attitudinal and information processing perspectives. Both focus on the cognitive states and/or processes that intervene between or mediate exposure to persuasive communications and changes in attitude, behavioral intention, or overt behavior. These two approaches, broadly termed the cognitive structure and cognitive response models, provide the theoretical foundations for this study.

The Cognitive Structure Model

The cognitive structure model has its theoretical roots in learning theory and emphasizes beliefs as the fundamental cognitive element (see Lutz & Swasy, 1977; Olson & Mitchell, 1975). For example, Fishbein proposed expectancy-value models that depict the causal relationships between beliefs and attitudes, intentions, and eventually behavior (cf. Fishbein & Ajzen, 1975). Specifically, attitudes are thought to be based on beliefs

(Fishbein, 1967):

$$A = \sum_{i=1}^n b_i e_i \quad (1)$$

where:

A = attitude toward (evaluation of) an object, action or event.

b_i = strength of belief about the "ith" attribute or consequence.

e_i = evaluation of the "ith" attribute or consequence.

n = number of salient beliefs.

By this model, Fishbein proposed that attitudes are a function of belief strength and the evaluative aspects of beliefs combined in an additive, compensatory manner. An extension of this attitude model (see Fishbein & Ajzen, 1975) relates the attitude concept to behavioral intentions (BI) which, in turn, are causally related to behavior (B) itself, as indicated by arrows in the following simplified schema:

$$\sum b_i e_i \rightarrow A \rightarrow BI \rightarrow B \quad (2)$$

Fishbein's model of attitude formation implies that a persuasive communication will affect attitudes indirectly via the processing of information contained in the communication which first creates b_i and e_i cognitive elements (see Fishbein & Ajzen, 1975). This approach to monitoring communication effects represents a step beyond the traditional research paradigm by actually measuring the cognitive variables (beliefs) that presumably mediate and determine attitude change and is becoming increasingly evident in the consumer behavior literature (cf. Kuehl & Dyer, 1977; Lutz, 1975; Mazis & Adkinson, 1976; Olson & Dover, 1976, in press). Beliefs formed by processing the information contained in the persuasive communication presumably are integrated into a pre-existing belief framework. According to Fishbein and Ajzen (1975), the resulting change in overall belief structure due to the integration of previous beliefs and beliefs derived from the communication information is the fundamental basis for attitude change and behavior modification.

Cognitive Response Models

An alternative conceptual view of communication impact is provided by cognitive response models which were developed explicitly for monitoring communication effects. The basic premise of these models is that the spontaneous unstructured cognitive responses (thoughts) elicited by a communication message act as mediators of attitude formation or change. In this view, a precise understanding of prediction of communication impact is not possible without attention to the cognitive thought responses emitted by the receiver upon exposure to the

¹This research was supported by a grant to the senior author from the Center for Research, College of Business Administration, The Pennsylvania State University.

²Associate Professor, Assistant Professor and Assistant Professor, respectively. This research was conducted while the second and third authors were doctoral students in Marketing at the Pennsylvania State University.

communication.

The measurement paradigm for a cognitive response study typically includes the following procedures (see Wright, 1974, for a more complete discussion). First, the subject is asked to either write down or verbalize all of his/her thoughts either during or immediately after exposure to a persuasive communication. These protocols are assumed to indicate the internal subvocal cognitive responses which occurred during and after exposure to the communication. Then, the cognitive responses are categorized according to various criteria, either by the subject or by the experimenter. Possible criteria include whether there is a "critical" or "favorable" orientation to the message, the perceived origin of the thought, and whether or not the response meets the prescribed definition for a class of cognitive responses. Examples of cognitive response categories include counterargument, source derogation, support arguments, and curiosity statements (see Wright, 1973, 1974). The third step is to compute cognitive response indices or scores based upon the categorization. These scores may represent either a simple addition of the number of counterarguments, source derogations, or support arguments, or more elaborately, use of a model in which each cognitive response type is weighted by subjective indications of importance (e.g., Wright, 1973). In both cases the scores represent an operationalization of the cognitive response mediators and are typically used to predict some criterion, usually attitude or intention.

The application of direct thought monitoring in communication research is in its infancy. Only a few studies, mostly in the social psychology literature, have used this conceptual and methodological orientation. In these investigations, the research emphasis has centered on the study of distraction effects (Osterhouse and Brock, 1970; Keating and Brock, 1974; Insko, Turnbull, and Yandell, 1974; Petty, 1977; Petty, Wells, and Brock, 1976), attitude change/persuasive impact (Cook, 1969, Wright, 1973, Petty, 1977), product versus task involvement (Krugman, 1967; Wright, 1973) repetition and media modality effects (Ray, Sawyer, and Strong, 1971; Wright, 1973; Ray, Ward, and Reed, 1976; Ray and Webb, 1976), source credibility (Cook, 1969) and fear arousing messages (Janis and Terwillinger, 1962). Wright (in press) has reviewed the current literature on cognitive response models. The potential of the cognitive response/thought monitoring approach to advertising research, although largely unexplored, appears strong and warrants the attention of consumer behaviorists.

A Combined Cognitive Structure/Cognitive Response Model

Both of the previously mentioned research orientations possess advantages and disadvantages as a modeling perspective for studying the impact of persuasive communications. However, if the two approaches are combined, it may be shown that each complements the other's weak points. Whereas the cognitive structure model focuses on the structural aspects of stored knowledge in the form of beliefs, the cognitive response model is more concerned with cognitive processing activity. Together, the two models can provide a relatively comprehensive model of communication impact. That is, cognitive responses (subvocal thoughts) can be seen as directly mediating the formation of new beliefs or changes in pre-existing beliefs. The belief structure so created would, in turn, mediate attitude formation and change. The advantages of, and potential for, this joint model approach were independently recognized and clearly presented in a particularly thorough paper by Lutz and Swasy (1977).

The present paper represents an initial attempt to examine the conceptual and predictive utility of the cognitive response perspective as a mediating process occurring between exposure to a communication and the formation of or change in cognitive structure variables such

as beliefs, attitudes and intentions. Due to space limitations, emphasis is on the more general and heuristic effects of cognitive responses on cognitive structure, as a function of exposure to advertising communications. Our long-range goal is to encourage and contribute to the development of a more precise and complete theory of communication effectiveness, at an individual level.

Methods

Research Design

To more clearly identify the mediating effects of cognitive responses, it was necessary to experimentally create alternative communications (in the form of simple, poster-like advertisements) that should lead to variations in magnitude and type of cognitive response. These differences, in turn, would be expected to have a variable impact on cognitive structure variables. Two factors were selected for manipulation, a product attribute and an ad content variable. Ads were created which communicated one of five arbitrary price levels for the ball point pen product--29¢, 49¢, 79¢, \$1.49, and \$3.95. Additionally, information regarding other salient (non-price) attributes of the pen was either present or absent. Thus the design was a 5 x 2, between-groups factorial involving 5 levels of price and the presence or absence of additional, non-price information about the product (see Table 1).

TABLE 1
Experimental Design

Information Contained in the Ad	Price Level				
	29¢	49¢	79¢	\$1.49	\$3.95
Price, only	n=15	n=17	n=17	n=16	n=17
Price & Additional Attribute Information	n=15	n=17	n=17	n=17	n=17

Basically, consumer subjects were exposed to one of the 10 advertisements (treatment combinations). Then their cognitive responses (thoughts) to the ad were recorded, followed by an extensive measurement of cognitive structure elements.

Although highly specific hypotheses were not considered possible given the current state of theory, we intuitively anticipated greater counterargument for higher prices. However, it was not clear a priori whether additional non-price information would lead to greater or less counterargument compared to price-only conditions. Moreover, we generally expected counterargument and support argument to mediate the formation of less positive and more positive cognitive elements, respectively.

Product and Subjects

Several factors influenced the choice of ball point ink as the focal product in this research. First, pilot interviews indicated a relatively simple cognitive structure for ball point pens in terms of salient product attributes. Thus, measurement complexity was kept within reasonable limits. Second, a relatively new, inexpensive pen possessing several unusual features had been recently introduced in the local market and was still unfamiliar to nearly all subjects. The actual brand name of this pen was used throughout the study, but is omitted in this paper.

Because ball point pens have a high usage rate among college students, a student population was considered appropriate for this research. Subjects were 165 undergraduate students from introductory business courses who were paid \$2.00 for their voluntary one-hour participation.

Advertisements

In all, ten separate ads were professionally created to correspond to the ten treatment combinations in the research design (see Table 1). Centered at the top of each in large type was the brand name followed by the headline, "Introducing the New ----- Ball Point Pen." In the center of each ad was a black and white drawing of the pen.

The experimental manipulations of price level and other information were accomplished by varying the remaining information in the ads. The price level was presented in the lower right corner of each ad -- e.g., "only 29¢," or "only \$3.95."³ In the price only conditions, no information other than price was given. In the "additional information" conditions, four other product features were listed in the lower right corner and very briefly described. These included (a) the physical construction of the ball point tip, (b) the tungsten carbide ball and thin, free flowing ink, (c) the attractive design, and (d) the comfortableness of using the pen.

Procedures

Subjects were run in groups corresponding to each treatment combination. The procedure by which subjects signed up to participate at a particular time was considered to be essentially equivalent to a formal randomization procedure.

Upon entering the research room, subjects were told in general terms that the study concerned their reactions to a new ball point pen and some promotional material for the product. Then, to meet University regulations, they signed an informed consent form certifying their voluntary decision to participate. Following this, subjects completed a two-page questionnaire regarding general knowledge of and purchase behavior towards ball point pens and the perceived importance of several pen attributes.

Next, subjects were told that a rough mock-up of a forthcoming ad for the new ball point pen had been obtained. This ad was described as one that might be used as a point-of-purchase poster, perhaps on a store counter-top. Each subject was then given a black and white copy of the ad to examine. After two minutes, subjects were instructed to turn the ad face down on the table and to write (on a blank sheet of paper before them) "all the things that came to mind as you read the ad." They were given exactly two minutes to record their cognitive responses. Following, subjects turned to the next page and began responding to the post-exposure questionnaire. To separate somewhat the measurement of cognitive response and cognitive structure, the first page of the questionnaire contained six structured and two open-ended questions regarding subjects' general reactions to the ad itself. Following this were questions measuring b_i , e_i , A_0 , A_{act} and BI elements of cognitive structure regarding the pen.

After completion of the questionnaire, the study continued with an actual trial experience with the pen and subsequent re-measurement of the cognitive structure variables (this post-trial data is not relevant for the present paper). At the conclusion of the study, subjects were debriefed (the deceptions used were thoroughly explained and justifications were given) and were paid and dismissed.

Dependent Variables

The elements of cognitive structure measured in this paper are those identified most closely with structural models of attitudes (cf. Fishbein and Ajzen, 1975) and

are represented in terms of their general causal relationships in (2).

The measurements of the belief strength (b_i) and evaluative aspects (e_i) were based on the vector model and methodological procedures proposed by Ahtola (1975). A major contribution of Ahtola's modification of the familiar Fishbein expectancy-value model (1) is the emphasis on a more specific and precise identification of the belief concepts. In the Ahtola view, a product attribute is conceptualized as a dimension containing several discriminable levels or amounts of the attribute. In perceiving or encoding information about a product attribute, a consumer essentially assigns the product to one or more of the discriminable categories along that attribute dimension. For example, a wine could be classified as sweet, slightly sweet, medium dry, dry, or very dry along the sweet-dry attribute dimension. From this perspective, each category or level represents a discriminable belief. If, for example, one is positively certain that a particular wine is medium dry, then the belief strength (b_i) for that category should be maximal and the b_i 's for the other categories should be zero. This basic notion suggests that people learn these categories (cf. Wyer, 1974) and thus cognitive structure research should focus on identifying the discriminable categories used by subjects to encode the meaning of a particular stimulus such as a brand.

Through pilot interviews with several undergraduate students, seven modally salient attributes for ball point pens were identified: price, ink flow consistency, comfort, appearance, ease of writing, writing quality, and value for the money. Discriminable categories (levels or amounts) for each attribute were developed based partially on experimenter intuition. For the categories along each attribute (e.g., very good, good, average, poor, and very poor value for the money), subjects indicated their belief strength by assigning 10 points to the several levels or categories such that the number of points indicated the strength of belief that the pen possessed each attribute level. Subjects also rated their evaluation of each attribute level.

An index of the overall evaluation associated with each attribute can be obtained by summing the $b_i e_i$ products over the category levels to yield a vector score of evaluation ($\sum b_i e_i$). An index of overall evaluation associated with the entire cognitive structure is represented by summing the separate vector/attribute scores ($\sum \sum b_i e_i$). Attributes toward the pen (A_0) and the act of buying the pen (A_{act}) were each measured by averaging the scores on the three 5-point scales following Fishbein's recommendations (1967). Intention to purchase for one's own personal use on the next shopping trip (BI) was measured on a 5-point scale (not likely--very likely to buy, 1-5, respectively).

Classifying the Cognitive Response Measures

The measures of subjects' cognitive responses to the various ads were classified by following closely the criteria proposed by Wright (1973). The two cognitive responses of major interest were counterargument and support argument. Briefly, counterargument is indicated by statements that are directed against (disagree or clash with) the product in general, its use, or a specific claim or idea contained in the ad. Support arguments are indicated by thoughts in favor of (agree with) the product, its use, or a specific claim or idea in the ad. Several other types of cognitive response were also identified, including: (a) curiosity statements -- general thoughts evidencing curiosity regarding product characteristics, (b) positive ad-related statements -- thoughts about positive aspects of the ad itself, and (c) negative ad-related statements.

³The actual price of the pen was 79¢ in most stores.

Treatment Effects on Cognitive Responses

Although this paper is primarily concerned with support and counterarguments, Table 2 presents, for the reader's interest, a summary of all the types of cognitive responses elicited by the two types of advertised information (collapsed across all price level conditions) and for the entire design. The overall average of 2.94 cognitive responses per subject is somewhat lower than that typically obtained in other studies (e.g., 3.8 to 10 or more). However, most of these other studies used much more complex communications with rather extensive copy thus providing more material to cognitively respond to. Moreover, most of these studies did not place an explicit time limit on the thought monitoring task and subjects may have generated items in the course of thought recording (cf. Wright, 1977a, 1977b, for excellent reviews and analyses of the methodology used in this research).

TABLE 2
Summary of All Cognitive Response Types

Type of Cognitive Response	Type or Information		Grand Total
	Price only	Price & Attributes	
Counterargument	53 (.65)	73 (.88)	126 (.76)
Support argument	38 (.46)	62 (.75)	100 (.61)
Curiosity statement	68 (.83)	32 (.39)	100 (.61)
Neutral statement	36 (.44)	31 (.37)	67 (.41)
Positive ad-related	7 (.09)	8 (.10)	15 (.09)
Negative ad-related	41 (.50)	36 (.43)	77 (.47)
Total	243(2.96)	242(2.92)	485(2.94)

Note: Mean number of responses are in parentheses

Note in Table 2 that curiosity thoughts dominate in the price-only ads where virtually no information other than price was available. In contrast, for the conditions involving information about price and other attributes, counter and support arguments were the most frequently occurring cognitive responses. Because these types of responses are thought to have the major mediating influence on cognitive structure, the remainder of the paper focuses only on support and counterargument effects.

Table 3 summarizes the effects of the experimental manipulations on the number of counterarguments and support arguments elicited by ad exposure. As expected, Price Level had a significant main effect on the number of both counterarguments and support arguments. Figure 1 illustrates this main effect -- i.e., the mean number of counterarguments as a function of Price. A Newman-Kuels analysis revealed that the 29¢ and \$3.95 prices produced significantly fewer and more counterarguments (p 's < .05) than the three intermediate prices, which did not differ from one another.

The experimental treatment effects on number of support arguments are less clear-cut. Although a main effect of Price Level was obtained (graphed in Figure 1), Type of Information was found to interact with Price. This two-way interaction is not readily interpretable and, therefore, is not illustrated. Finally, the significant main effect of Type of Information was such that more support arguments were elicited by price-plus-other-attribute information ($\bar{X} = .75$) than price-only information

($\bar{X} = .46, p < .05$).

TABLE 3
Effects of Price Level and Type of Advertised Information on Number of Counter and Support Arguments

Price Level	Number of Counterarguments		Number of Support Arguments	
	Price only	Price & Attributes	Price only	Price & Attributes
29¢	2 (.13)	7 (.47)	7 (.47)	19 (1.27)
49¢	10 (.59)	15 (.88)	8 (.47)	10 (.59)
79¢	8 (.47)	17 (1.00)	17 (1.00)	11 (.65)
\$1.49	13 (.81)	11 (.65)	5 (.31)	17 (1.00)
\$3.95	20 (1.18)	23 (1.35)	1 (.06)	5 (.29)

F-ratios from ANOVAs

Information Type (A)	F-ratio	F-ratio
Information Type (A)	2.59	5.20*
Price Level (B)	4.35**	4.02**
A x B	<1	2.76*

* $p < .05$; ** $p < .01$

Note: Means are given in parentheses

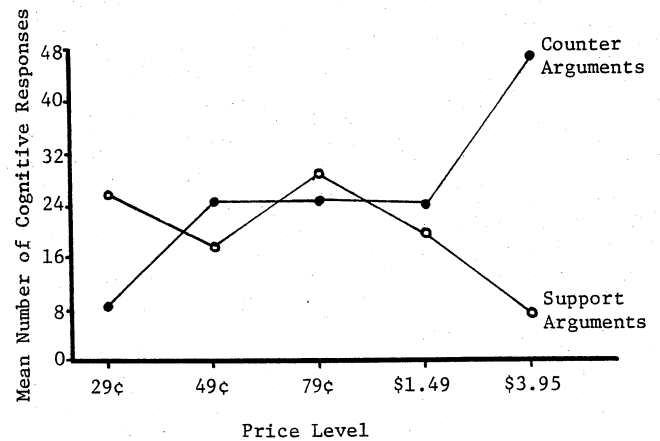


FIGURE 1: Main Effects of Price Level on Magnitude of Counter and Support Argument.

Given the above evidence that the experimental treatments did cause differences in type and level of cognitive response, our attention turns to the mediating effects of counter and support arguments on elements of cognitive structure. To keep this paper within its space limitations, it was necessary to examine only selected cognitive elements. In other papers, we will tackle the complex analyses involved in studying cognitive response effects on individual belief elements. In the present paper we will concentrate on the more global cognitive elements such as attitudes and intentions.

Cognitive Response Effects on Cognitive Structure

To determine the relationship between counter and support

arguments and these selected cognitive structure elements, two analyses were performed. The first involved collapsing the price level factor and computing correlations between number of counter and support arguments and selected cognitive variables for the two information type conditions. Table 4 summarizes the results of this analysis. As expected, counterargument was negatively related to cognitive structure variables, while support argument was positively related to cognitive structure elements.

TABLE 4
Correlations Between Number of Counter and Support Arguments and Selected Cognitive Elements

Cognitive Structure Variables	Cognitive Responses	
	Counter Arguments	Support Argument
Price Only Condition (n=82)		
$\Sigma b_i e_i^a$	-.41**	.24*
$\Sigma \Sigma b_i e_i$	-.21*	.16
A_o	-.07	.10
Aact	-.13	.03
BI	-.35**	.21*
Price & Attribute Condition (n=83)		
$\Sigma b_i e_i^a$	-.29**	.25*
$\Sigma \Sigma b_i e_i$	-.31**	.38**
A_o	-.27**	.36**
Aact	-.23*	.26**
BI	-.35**	.39**
All Respondents (n=165)		
$\Sigma b_i e_i^a$	-.31**	.26**
$\Sigma \Sigma b_i e_i$	-.23**	.30**
A_o	-.16*	.30**
Aact	-.17*	.19**
BI	-.33*	.32**

^aVector evaluation score for "value for money"

* p < .05

**p < .01

Both counterargument and support argument were significantly related to the belief vector regarding value for the money ($\Sigma b_i e_i$), the overall cognitive structure index ($\Sigma \Sigma b_i e_i$), and intentions to purchase. The relationships with attitude (A_o and Aact) were generally significant but somewhat weaker. Moreover, the relationships between cognitive response and structural variables are generally stronger and more consistent for the condition involving information about price and other attributes than for the price-only ads. When the information-type factor is collapsed, both counter and support argument responses are significantly related to all the more global elements of cognitive structure, although none of the relationships is particularly strong (r 's from .16 to .33, p 's < .05).

An alternative and perhaps more useful perspective on cognitive response influence is gained by examining the magnitude of the mediating effect on selected cognitive variables. This analysis was accomplished by collapsing the price level factor and dividing the subjects in each information type group into those making no counterarguments ($CA = 0$) vs. those with some ($CA \geq 1$) counterarguments ($X = 1.47$). In effect, this internal analysis created a 2 x 2 design composed of two levels of information type and two levels of counterargument magnitude (some and none). The cell means and ANOVA results for several cognitive variables are presented in Table 5. Scores for Aact are not shown because no effects were found for this variable. Note that Counterargument was consistently and significantly related to the value-for-the-money belief vector, and to belief structure, product attitude, and purchase intention. Information Type had generally weaker main effects. Only one (relatively weak) interaction was obtained. In all cases, the effect of counterargument was to lower the positiveness (increase the negativity) or weaken the cognitive variable.

The same internal analysis was conducted by splitting the sample based on those subjects who reported no support arguments (n=98) and those who reported one or more support arguments (n=67). When this blocking factor was combined with the information-type treatment conditions, a 2 x 2 internal analysis was again conducted on the selected cognitive variables of interest. The results are presented in Table 6. Support Argument was significantly related (main effects) to each of the cognitive elements, such that the variable was more positive or stronger when subjects engaged in support arguing. Information Type had some main effects on the belief and attitude variables. No interaction effects were obtained.

Conclusions and Implications

The results of this study provide support for the notion that cognitive responses to persuasive communications (ads in this case) mediate the effect of the message on elements of cognitive structure (Lutz and Swasy, 1977; Wright, 1973; 1974). Both counterarguments and support arguments were found to be related to a wide range of cognitive variables including beliefs, attitudes, and purchase intentions. Although strict cause and effect statements cannot be supported by this design, it seems logically reasonable to presume that cognitive responses precede and thus influence the formation of cognitive structure elements. The direction of influence of the types of cognitive responses was, of course, opposite, with counterarguments leading to less positive or weaker cognitive states and support arguments leading to more positive or stronger cognitive states. The broad consistency of these effects across different analyses and over a variety of cognitive measures, lends additional credence to the basic conceptual ideas about cognitive responses and to their construct and measurement validity. In sum, cognitive response theory appears to have considerable utility for improving our generally weak theories of the persuasive communications process, and therefore warrants specific attention in future advertising/communication research.

The relatively strong and consistent mediating effects of support and counterargument responses as evidenced in this study may be seen as even more robust when contrasted with earlier research. The bulk of the past research has used rather strong, overtly persuasive communications--ones quite likely to maximize cognitive responses, especially support and counterarguments (Wright, 1977a, 1977b, in press). In contrast, the present study examined ads which were primarily informational in "flavor" rather than distinctly persuasive. This was particularly true of the price-only ads. One point is that consumers may counterargue and/or support argue

TABLE 5
Effects of Counterargument and Type of Advertised
Information on Selected Elements of Cognitive Structure

	Selected Cognitive Elements							
	$\Sigma b_{ii} e_i^a$		$\Sigma \Sigma b_{ii} e_i$		A_o		BI	
	CA=0	CA>1	CA=0	CA>1	CA=0	CA>1	CA=0	CA>1
Price only	11.74	1.39	69.43	50.56	3.38	3.31	2.93	2.22
Price, plus other attributes	17.20	7.84	98.05	63.65	3.92	3.45	3.28	2.44
F-ratios from ANOVAs								
Information Type (A)	5.45*		6.94**		12.13**		1.36	
Counterargument (B)	18.02**		12.57**		6.63*		15.62***	
A x B	1.78		2.71		6.29*		<1	

*p < .05, **p < .01, ***p < .001

^aVector score for value-for-the-money attribute

TABLE 6
Effects of Support Arguments and Type of Advertised
Information on Selected Elements of Cognitive Structure

	Selected Cognitive Elements									
	$\Sigma b_{ii} e_i^a$		$\Sigma \Sigma b_{ii} e_i$		A_{act}		A_o		BI	
	SA=0	SA>1	SA=0	SA>1	SA=0	SA>1	SA=0	SA>1	SA=0	SA>1
Price only	5.09	11.25	57.43	69.14	3.47	3.44	3.32	3.40	2.51	2.89
Price, plus other attributes	9.25	15.85	67.09	94.49	3.48	3.82	3.42	3.97	2.55	3.18
F-ratios from ANOVAs										
Information Type (A)	5.08*		6.65*		3.70		12.43***		1.29	
Support Argument (B)	8.95**		8.71**		4.00*		14.66***		7.65**	
A x B	<1		<1		2.24		2.56		<1	

*p < .05, **p < .01, ***p < .001

^aVector score for value-for-the-money attribute

with message content that does not directly attack established beliefs. Another is that such processes may occur and mediate impact on cognitive structure even for seemingly unimportant, low involvement products such as ink pens. A third and very important applied application of the present results is that specific aspects of ad content (such as the type of information conveyed) may have a fairly strong effect on the cognitive responses produced and, in turn, on subsequent effects on cognitive structure. This latter idea suggests that advertising research could focus on determining specific

ad content variables that would minimize counterarguments (or maximize support argument), thus "paving the way" for cognitive structure effects favorable to the firm.

In conclusion, it would seem that cognitive response theory and methodology warrant the attention of consumer researchers. When combined with other theoretical ideas regarding information processing, this approach may offer new insights into the communication process. Moreover, the entire research area holds promise of applicability to real world communication problems.

References

- Olli T. Ahtola, "The Vector Model of Preferences: An Alternative to the Fishbein Model," Journal of Marketing Research, 12(February, 1975), 52-59.
- Bobby J. Calder, Chester A. Insko, and Ben Yandell, "The Relation of Cognitive and Memorial Processes to Persuasion in a Simulated Jury Trial," Journal of Applied Social Psychology, 4(1974), 62-93.
- Thomas D. Cook, "Competence, Counterarguing, and Attitude Change," Journal of Personality, 37(1969), 342-358.
- Martin Fishbein, "A Consideration of Beliefs and Their Role in Attitude Measurement," in M. Fishbein (ed.), Readings in Attitude Theory and Measurement, (New York: Wiley, 1967), 257-266.
- Martin Fishbein and Icek Ajzen, Belief Attitude, Intention and Behavior, (Reading, Massachusetts: Addison-Wesley, 1975).
- Chester A. Insko, William Turnbull, and Ben Yandell, "Facilitative and Inhibiting Effects of Distraction on Attitude Change," Sociometry, 37(1974), 508-528.
- Irving L. Janis and Robert F. Terwilliger, "An Experimental Study of Psychological Resistances to Fear Arousing Communications," Journal of Abnormal and Social Psychology, 65(1962), 403-410.
- Herbert E. Krugman, "The Measurement of Advertising Involvement," Public Opinion Quarterly, 30(1967), 583-596.
- Philip J. Kuehl and Robert F. Dyer, "Application of the 'Normative Belief' Technique for Measuring the Effectiveness of Deception and Corrective Advertisements," in W. Perreault, Jr. (ed.), Advances in Consumer Research: Vol IV, (Atlanta: Association for Consumer Research, 1977), 204-212.
- Richard J. Lutz, "Changing Brand Attitudes Through Modification of Cognitive Structure," Journal of Consumer Research, 4(March, 1975), 49-59.
- Richard J. Lutz and John L. Swasy, "Integrating Cognitive Structure and Cognitive Response Approaches to Monitoring Communications Effects," in W. Perreault, Jr. (ed.), Advances in Consumer Research: Vol IV, (Atlanta: Association for Consumer Research, 1977), 363-371.
- Michael B. Mazis and Janice E. Adkinson, "An Experimental Evaluation of a Proposed Corrective Advertising Remedy," Journal of Marketing Research, 13(May, 1976), 178-183.
- Andrew A. Mitchell and Jerry C. Olson, "Cognitive Effects of Advertising Repetition," in W. Perreault, Jr. (ed.), Advances in Consumer Research: Vol. IV, (Atlanta: Association for Consumer Research, 1977), 213-220.
- Jerry C. Olson and Philip Dover, "Effects of Expectation Creation and Disconfirmation on Belief Elements of Cognitive Structure," in B. B. Anderson (ed.), Advances in Consumer Research: Vol. III, (Cincinnati: Association for Consumer Research, 1976), 168-175.
- Jerry C. Olson and Philip A. Dover, "Cognitive Effects of Deceptive Advertising," Journal of Marketing Research, in press.
- Jerry C. Olson and Andrew A. Mitchell, "The Process of Attitude Acquisition: The Value of a Developmental Approach to Consumer Attitude Research," in M. J. Schlinger (ed.), Advances in Consumer Research: Vol. II, (Chicago: Association for Consumer Research, 1975), 249-264.
- Robert A. Osterhouse and Timothy C. Brock, "Distraction Increases Yielding to Propoganda by Inhibiting Counterarguing," Journal of Personality and Social Psychology, 15(1970), 344-358.
- Richard E. Petty, "The Importance of Cognitive Responses in Persuasion," in W. Perreault, Jr. (ed.), Advances in Consumer Research: Vol. IV, (Atlanta: Association for Consumer Research, 1977), 358-362.
- Richard E. Petty, Gary L. Wells, and Timothy C. Brock, "Distraction Can Enhance or Reduce Yielding to Propaganda: Thought Disruption Versus Effort Justification," Journal of Personality and Social Psychology, 34(September, 1976), 874-884.
- Michael L. Ray, Alan G. Sawyer, and Edward C. Strong, "Frequency Effects Revisited," Journal of Advertising Research, 11(February, 1971), 14-20.
- Michael L. Ray, Scott Ward, and Jerome B. Reed, "Pre-testing of Anti-drug Abuse Education and Information Campaigns: Summary Report of a Marketing Science Institute Special Project," in R. Ostman (ed.), Communication Research and Drug Education, (Beverly Hills, California: Sage Publications, 1976).
- Michael L. Ray and Peter Webb, "Experimental Research on the Effects of TV Clutter: Dealing With the Difficult Media Environment," Report No. 76-102, Marketing Science Institute, Cambridge, Massachusetts, (April, 1976).
- Peter L. Wright, "The Cognitive Processes Mediating Acceptance of Advertising," Journal of Marketing Research, 10(February, 1973), 53-62.
- Peter L. Wright, "On the Direct Monitoring of Cognitive Response to Advertising," in G. D. Hughes and M. L. Ray (eds.), Buyer/Consumer Information Processing, (Chapel Hill: University of North Carolina Press, 1974), 220-248.
- Peter L. Wright, "Research on Ad-Stimulated Thought Processes," Unpublished working paper, Graduate School of Management, Stanford University, February, 1977(a).
- Peter L. Wright, "Speculations and Evidence About Audience Responses to Mass Media Advertisements," a paper presented at the conference, "The Effect of Information on Consumer and Market Behavior," Carnegie-Mellon University, May 20 and 21, 1977(b).
- Peter L. Wright, "Cognitive Responses to Mass Media Advocacy and Cognitive Choice Processes," in R. Petty, T. Ostrom, and T. Brock (eds.), Cognitive Response to Communication Processing Approach, (New York: Wiley, 1976).
- Robert S. Wyer, Jr., Cognitive Organization and Change: An Information Processing Approach, (New York: Wiley, 1974).

THE ORGANIZATION OF PRODUCT INFORMATION IN MEMORY IDENTIFIED BY RECALL TIMES¹

Eric J. Johnson (student), Carnegie-Mellon University²
J. Edward Russo, University of Chicago

Abstract

A chronometric technique is used to examine whether consumers' memory structure for product information reflects the input organization of that information, or whether people will reorganize into a preferred structure, either brand-based or attribute-based. The evidence clearly favors congruence between memory structure and input organization, although some reorganization did occur.

This report has two purposes: (1) to call attention to the role played by memory for product information in purchase decisions, and (2) to introduce a chronometric technique to the study of the structure of remembered product information. When consumers make purchases, they use product information from two sources. The external environment contains package labels, point-of-purchase displays, posted prices, etc., relevant to the purchase decision. However, most information needed for the purchase decision is retrieved from memory. Information like use experience, family preferences, word-of-mouth communications, and advertising (except point-of-purchase advertising) must all be mediated by memory. Further, if all brands are not available from a single vendor, then information about the missing ones must come solely from memory. The point is that product information retrieved from memory plays an essential role in almost all purchase decisions. Unfortunately, research into consumer decision making has largely ignored this role (Olson, 1977).

Role of Memory

A processing history of product information can be outlined as follows:

1. encoding of product information from external sources.
2. retention of product information in memory.
3. retrieval and use of product information in the purchase decision process.

The second stage, retention in memory, is the focus of this paper. It will be useful, however, to consider the interactions between retention and the other two stages.

The Effect of Retention on Decision Strategies

Does the nature of the stored information determine the type of decision process? For example, if product information is organized around brand names, as many researchers have suggested, will the decision rule have to be compatible with this information structure? Two studies suggest that the answer may be yes. Simon and Hayes (1976) used different cover stories for the same problem to induce different internal representations of the problem information. The subjects then selected problem solving strategies that were compatible with

¹The authors would like to thank John Herstein for introducing them to the problem, Patrick Langley for helpful discussions, and Jeff Godliss for running subjects.

²Order of authorship is alphabetical; the authors' contributions are equal.

their specific internal representations of the information. Different representations led to different strategies. If purchase decisions are seen as a type of problem, the Simon and Hayes result implies that different representations of this same product information may well lead to different decision strategies. Further, this implication should hold whether the representations are internal or external.

The last assertion has been explicitly confirmed by Bettman and Kakkar (1977), except that they relied entirely on externally available rather than remembered information. Each of their subjects was shown product information in one of three formats: a brand-organized display, in which information was available most easily by brand name; an attribute-organized display, in which the same information was available most easily by attribute; and a matrix display, which equated the availability of the information. Although consumers could process the available information however they wished, Bettman and Kakkar found that the decision rule was usually chosen to be compatible with the format of the product information. Thus, decision rules that required the information to be processed by brand were most frequent with the brand-organized format. This study provides direct evidence that consumer decision rules are flexible and will depend on the structure of the available product information.

The Effect of Encoding on Retention

Just as the structure of product information can influence the subsequent decision strategies, the process used in encoding the information can determine how the information is stored. For example, if product information is presented in a brand-organized fashion, will it be remembered in a brand-organized structure? Alternatively is there a preferred memory structure into which product information is always organized no matter what the original presentation and encoding format?

Unfortunately, the authors know of no studies that have investigated the relation between the encoding of product information and its organization in memory. The experiment presented in subsequent sections was designed to address this issue.

The Realworld versus the Laboratory

The existing research literature has largely ignored the role of memory in purchase decisions. Most laboratory studies attempt to bypass any memory stage. Instead, all product information is provided externally.

We believe that researchers in consumer behavior have not deliberately shunned memory issues *per se*. Rather, this omission follows naturally from the information acquisition paradigm, which dominates research on consumer choice behavior. Within this paradigm experiments are designed to rely as much as possible on externally provided information because the information acquisition sequence provides the necessary data base. Since acquisition from memory is unobservable, it can provide no insight into the purchase process.

This information acquisition paradigm is, of course, unrepresentative of realworld purchases. Most such purchases rely primarily on remembered information.

Thus, an unfortunate consequence of a generally successful research paradigm has been a systematic gap in our investigation of consumer behavior.

Representation of Knowledge

Two alternative structures have been proposed to represent a consumer's product knowledge. This knowledge can be stored by brand or by attribute (Bettman, 1978; Payne, 1976). Storage by brand assumes that a person's knowledge about a product class is organized around each brand. To take peanut butter as an example, a verbal statement of brand-organized knowledge might be: "Skippy tastes very good, contains preservatives and costs \$.99 for an 18 oz. jar." In contrast, attribute-organized knowledge could be verbalized as: "Price, let's see, Skippy costs \$.99, Jif costs \$.99, but Peter Pan costs only \$.95." Note that whether the product information is stored by brand or by attribute, the information itself is identical. The price of Skippy is stored as \$.99 in either structure. The concern is not with the content of memory but with its structure or organization.

This conceptualization of product knowledge should be distinguished from such theories as those underlying multidimensional scaling and semantic differential scales. The emphasis of the present research is not on the content of the stored information, but with the organization of the information itself. Although not everyone will share the same sets of brands and attributes, especially across different products, it is suggested that preferred organization is independent of content.

Representations in Cognitive Psychology

A major concern of cognitive psychology is the representation of knowledge. There are several highly related formalisms for representing the structure of long-term memory, including list structures (Simon and Newell, 1974) and semantic networks (Anderson, 1976). References to these and other readings in the representation of knowledge can be found in Norman (1976, pp.197-198). Because list structures can be most easily simplified, Figure 1 presents the previous example in a list structure format. Both the brand-organized and attribute-organized versions are shown.

FIGURE 1
List Structure Representations of Product Knowledge

	(Skippy	(Taste	Very Good)
		(Preservatives	Yes)
A		(Price	\$.99)
		(Size	18 oz.)
	(Jif	(Taste....	
	(Price	(Skippy	\$.99)
B		(Jif	\$.99)
		(Peter Pan	\$.95))
	(Taste	(Skippy....	

The top structure (A) is brand-organized, while below it (B) is an attribute-organized list structure.

The essence of the list structure representation is the use of the "next" relation to connect items. Thus, in Figure 1A Skippy is followed by a list of its attributes, whereas in Figure 1B, Skippy is followed by the other brands. The links created by the next relation can be traversed to get from any item in the complete list structure to any other item. Of course, the nature of the list structure can either facilitate or hinder such activity. In the brand-organized list

(Figure 1A), if one is at Skippy-Preservatives-Yes, it is easier to reach Skippy-Size 18 oz. than to reach Jif-Preservatives-Yes (not shown).

The Response Time Technique

The verification of structural representations such as those depicted in Figure 1 is often accomplished by using chronometric techniques. In essence, these techniques assume a direct relation between the amount of processing and the time taken to perform a task. If the task is recall, the time to recall is assumed to measure the "distance" in the memory structure between the recall cue and the target item.

It is important to realize that it is not just access time that makes up time to recall. The subject may, in fact, have the needed information in memory, but in a form different from that necessary for the correct response. Therefore, some time must be spent converting the internal representation to a more useful form. However, whether more cognitive processing is required by greater "distance" or by the conversion of "near" information, the longer recall time accurately indicates the greater effort needed to recall the given item.

Chronometric analysis has been developed into a sophisticated methodology. Although only the basic principles will be used in this report, it is worth noting two elaborations. In most tasks, it is possible to trade speed for accuracy. Such a trade-off threatens the validity of recall times in identifying cognitive structure. In practice, this trade-off is avoided by enforcing minimal error rates through instruction and training. However, more powerful techniques are available for situations with large or unstable error rates or when the retrieval dynamics must be known in detail (e.g., Doshier, 1976). A second elaboration is the partitioning of the total response time into segments that measure the durations of several, nonoverlapping stages of processing. This version of chronometric analysis was developed by Sternberg (1969), and many successful applications could be cited (see Chase, 1977).

The analysis of retrieval times is one of a surprisingly large number of techniques for investigating memory representations (for a review, see Bower and Tulving, 1974). A more common technique is to ask the subject to respond to a presented item as the "same" as or "different" from some member of the memorized set of items. Both techniques require the recall of a remembered item. However, the verification task seems more removed from most purchase situations. Occasionally a consumer may search memory to verify a manufacturer's claim, but usually the consumer is recalling the relevant product knowledge as rapidly as possible. Thus, the recall time technique not only enables the discrimination between brand-organized and attribute-organized memory structures, but also relies on behavior that is similar to the retrieval of information in real shopping situations.

Experimental Rationale

The focus of our experiment is the relation between the first two stages of the "processing" of product information, namely encoding and retention. Specifically, we inquire whether the encoding format, brand-based or attribute-based, determines the storage organization or whether there is a systematic tendency to reorganize into a preferred memory structure regardless of input format.

The basic design had subjects learning a booklet of product information. The total information set was a complete brand x attribute matrix with four brands and four attributes. Each page of the booklet consisted of a column of information headed by either a brand name or an attribute name. The column itself contained, respectively, either the four attributes and their values for that brand or the four brands and their values for that attribute. These two input organizations were called brand input and attribute input, respectively.

The information in a complete brand x attribute matrix is rarely available in the realworld. However, it was important to the success of this experiment to treat brands and attributes equally. That is, any preferences for brand-based or for attribute-based memory structures had to be separable from encoding factors that favored one over the other. For this reason a complete, symmetric brand x attribute matrix was used. Future studies can address the effects of encoding factors, especially as they occur in realworld sources of product information.

After learning a booklet, the subjects performed a cued speeded recall task. They were presented with one of the brands or attributes and asked to recall, as quickly as possible, all the information associated with that cue. For example, if the cue was Skippy, a subject had to report each of the four attributes followed by its value for Skippy.

The cued speeded recall task was the same whether the information had been learned by brand or by attribute. The response times, however, should be different. In the brand input condition, the links between attributes within a brand should permit relatively fast recall of information to a brand cue. In contrast, a cue like price is separated from the necessary information by a larger set of links. Thus, our experiment depends on this simple principle: if information is stored by brand, it can be recalled more quickly by brand than by attribute; if it is stored by attribute, it should be recalled faster by attribute than by brand.

Method

Materials

Two brand-attribute matrices were used, one for air conditioners and one for cooking oils. These products were selected to represent both an infrequently purchased major durable and a common nondurable with a high repurchase rate. For air conditioners the brands were Admiral, Carrier, Frigidaire and Westinghouse, and the attributes were average price, circulation, energy use, and indoor noise. All 16 values took on one of three levels, high, medium or low. This matrix was adapted from evaluations of air conditioners published in Consumer Reports (July, 1973; and July, 1975). For cooking oils the four brands were Crisco, Kroger, Planters, and Wesson, and the four attributes were fat count, flavor, smoke point, and unit price. All 16 values took on one of three levels, fair, good, or excellent. This matrix was adapted from an evaluation of cooking oils published by Consumer Reports (September, 1973). Accompanying the attributes for each product category was a description of the meaning of the attribute. Subjects began the learning of product information by studying these attribute descriptions until they felt comfortable in their understanding of all of them.

The selection of particular brands and the phrasing of attribute labels were designed to maintain equal syllable counts for all rows and columns of a product matrix. This insured that response time was a measure of re-

trieval difficulty and not the output (speaking) time. Preliminary tests of the time needed to read each row and column within a matrix confirmed the approximate equality of the time to speak the eight files in a matrix.

Design

There were three main factors: A, input organization (brand or attribute); B, product (air conditioners or cooking oil); and C, probe type or cue type (a brand name or an attribute name). A fourth factor (D) is the specific cue, e.g., a specific one of the four brand names of air conditioners. This factor is nested within combinations of B and C. Each of the factors was fixed.

The three main factors were fully crossed (i.e., factorial) with one critical exception. Subjects could not be expected to learn information about the same product (Factor B) twice, once in each of the two input organizations (Factor A). Instead, each subject memorized both booklets once, alternating input formats (with order counter-balanced). The design for Factors A and B was a 2 x 2 Latin square fractional factorial design (see Kirk, 1968, p. 407). The main import of this design is the loss of all information about the A x B and A x B x C interactions. The only possible alternative design required a single subject to memorize both booklets in the same input organization. This split-plot design, by confounding group and input format effects, made the main effect of input (Factor A) indeterminate, which was even less desirable than losing an interaction effect.

The 2 x 2 Latin square divided the subjects into two groups (with 10 subjects in each group). Because 8 recall times (4 attribute cues and 4 brand cues) were obtained from both products and from each of the 20 subjects, a total of 320 observations formed the data base.

Subjects

Twenty-four undergraduates at Carnegie-Mellon University participated in the experiment. They were either volunteers (2) or were fulfilling a course requirement for introductory psychology (22). Two subjects were eliminated because they made more than three errors on the sixteen possible trials. Additionally, two subjects took an excessive amount of time to learn the information booklet for the first product. They were not asked to finish the experiment and their data were discarded. Thus, a total of twenty subjects provided data for the experiment.

When the response time data were analyzed, the means of 19 subjects ranged between 8.8 sec and 16.4 sec. The remaining subject's mean was 33.4 sec. Because this subject's responses were so unique and because they would thoroughly disrupt any pattern of results that held over the other 19 subjects, he was dropped from the study. One of the two subjects who failed to pass the error criterion had participated in the same condition as the dropped subject. This former subject was returned to the sample to maintain ten subjects per group.

Procedure

Subjects were told that they were participating in a study examining consumers' knowledge of various products. They were then asked to learn the contents of the first booklet. Instructions emphasized that the material in the booklet had to be remembered accurately. Subjects were also asked not to switch between pages of the booklet but rather to look at the book one page

at a time. After they had informed the experimenter that they thought they had learned the booklet, subjects were administered a criterion test. Subjects were asked to name the value of four different product-attribute combinations. If a subject could not successfully name the four values, he was given the booklet and told to try again. Subjects who passed the criterion test went on to the cued speeded recall task. The total time until the subject passed the test was recorded.

In the recall task subjects were presented either an attribute or a brand name. If the probe was an attribute name, they were told to recall all of the appropriate brand names and their values. For a brand name probe, they recalled the various attributes and values for that brand. The recall time was measured (in msec) from the onset of the probe to the end of the list. For each booklet the subject received four attribute probes and four brand probes. Order was randomized across subjects.

After completion of the first recall test, subjects were given the second booklet, this time with the opposite organization. The procedure for criterion testing and cued speeded recall was repeated.

Finally, subjects were presented an empty matrix for the first product class. The four attributes and the four brands were given but the sixteen values were missing. The subjects were asked to recall these values, and to write them down, in whatever order they wished. The output order was recorded. This was done for the second product as well. Then the subjects were debriefed and released.

The written free recall measure is not necessarily an independent measure of memory structure. Inherent in the cued memory task is recall by both brand and attribute. This task demand could cause reorganization which could well affect subsequent free recall. Also the time between learning the information and recall varies and the structure as therefore measured may be different than that measured with the cued recall. High convergence between these two recall methods may well wait until they are tried under a different experimental procedure.

Results

Encoding Time and Error Rate

A preliminary analysis tested for differences in the time necessary to learn the product information. An analysis of variance showed that learning times did not differ between the types of input organization, the different products or the two groups of subjects. Also, no interactions among these factors were significant. This precludes the possibility that differences in recall time could be due to differences in how well the lists were initially learned.

In the same way, differences in recall time could be confounded with differences in error rates across conditions. If speed is traded off for accuracy, then a difference in recall time can be due to different error rates as well as to the effect of an independent variable. Error rate was defined as the proportion of cued recalls for which at least one of the four required values (belonging to a single brand or a single attribute) was incorrect. The overall error rate was 7.8% (25 of 320). Using individual χ^2 -tests for the equality of two proportions, no significant differences in error rate were found between levels of input organization, product category, or cue type.

Taken together, these analyses imply that differences in error rates or thoroughness of learning can be ruled out as explanations for observed differences in recall speed.

Recall Times

Differences in recall times were tested by an analysis of variance. An examination of the residuals of the appropriate model indicated that a log transformation was needed to stabilize the variance over different levels of recall time. The results of the analysis of variance on the transformed recall times are reported in Table 1. The total degrees of freedom, 294, reflect the removal of the 25 recall trials on which an error occurred. Note that there are no replications in the design. The error term was estimated from the Group x Probe and Subject x Probe interactions. If these terms include any systematic efforts, the resulting F tests will be conservative.

TABLE 1
Analysis of Variance of Recall Times

Source	df	SS	MS	F
Between Subjects:				
Group (G) x	1	.01	.01	0.01
Subjects within G	18	10.82	.60	7.50***
Within Subjects:				
A	1	.00	.00	0.02
B	1	.83	.83	10.39**
C	1	.45	.45	5.59*
D (within B and C)	12	5.06	.42	5.26***
A x C	1	3.13	3.13	39.05***
B x C	1	.42	.42	5.28*
G x C	1	.65	.65	8.11**
Subjects x C	18	2.95	.16	2.05**
Subjects x A	18	3.78	.21	2.63***
Subjects x A x C	18	3.69	.20	2.56***
Subtotal	91	31.80	.35	
Error	203	16.27	.08	
Total	294	48.07		

* $p < .05$
** $p < .01$
*** $p < .001$

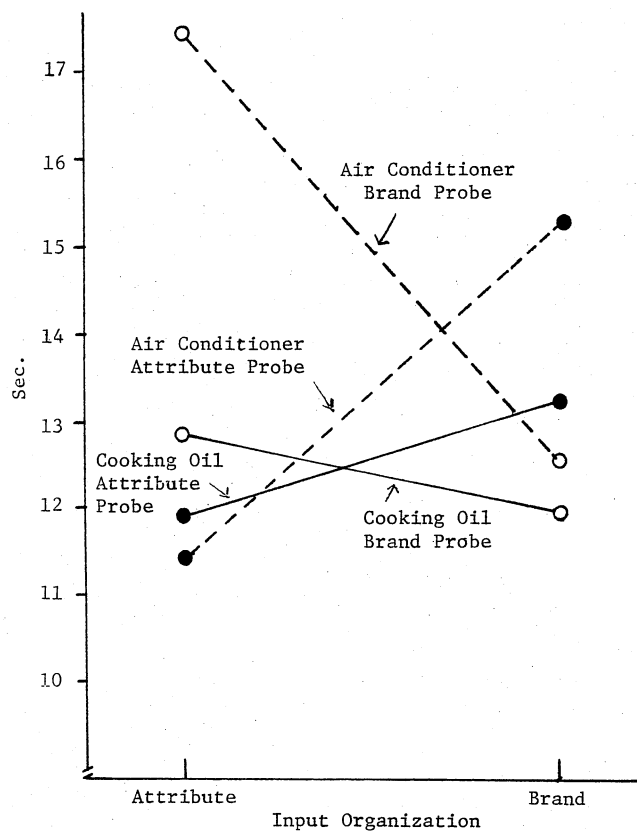
The critical prediction concerned the A x C interaction. This interaction captures the effect of storage organization (as induced from input organization) on the cued recall. It was predicted that recall would be faster when the probe type matched the structure stored in memory. Attribute probes would result in faster recall with attribute-organized information, and brand probes would be faster for brand-organized information. The "incompatible" probes would lead to slower recall.

As predicted, the A x C interaction is statistically significant and, furthermore, is the largest effect (in terms of mean squares) in the model. Before pursuing this and related effects, the subject and cue differences are noted. The presence of significant subject differences is typical in chronometric analysis. They are potentially attributable to many factors including motivation and strategy differences. There was no difference between groups, indicating that

subject differences were not systematically related to group. Note that in this design the A x B interaction is confounded with group. Therefore, the lack of a significant group effect validates the assumption of zero A x B interaction. The only other possibility is that both effects were of equal magnitude and opposite direction, thus cancelling each other out. Individual cue effects (Factor D) were also significant. This result was explored with a series of *a posteriori* contrasts. These tests revealed only one significant difference: for cooking oil the cue "cost" was faster than the cue "fat count." In general, the differences among cues exhibited no systematic relation, and they are of little intrinsic interest. The observed differences may be due to any articulation time or to irrelevant factors affecting memory, such as word frequency or concreteness.

To communicate the results involving the major independent variables, means over subjects and cues were computed. These values are plotted in Figure 2. The two pairs of crossing lines represent the A x C interaction, separately for the two product categories. Using *a priori* contrasts, both interactions were found to be significant. For air conditioners $F(1, 203) = 47.30, p < .001$; for cooking oils the interaction was less dramatic, $F(1, 203) = 5.93, p < .05$. Both the contrasts and an inspection of Figure 2 reveal that, though the predicted effect of compatibility between cue type and input organization is present for both product categories, the size of this effect differs across products.

FIGURE 2
Mean Recall Times



Returning to Table 1, it can be seen that both Factors B (product category) and C (cue type) show significant effects. As shown in Figure 2, cooking oil information was recalled faster than air conditioner information, holding all other factors constant. One possible explanation for this difference is articulation time. Though speaking times were approximately equalized across all brands and sizes within a product category, there were differences between categories. The total number of syllables over the eight brand or attribute names is 27 for air conditioners and 16 for cooking oil. Thus, the articulation time would be expected to be longer for air conditioners, as confirmed by the means in Figure 2. The other possibility is that cooking oil information is genuinely easier to recall. However, both products were learned as quickly and both yielded equal error rates during recall. These results suggest that the difference in recall times is not due to memory retrieval but rather to an output factor like articulation.

Differences between cue types are also apparent in Figure 2. The mean recall times were 12.97 sec for attribute probes and 13.70 sec for brand probes. This difference, though relatively small (2.7%), is intriguing. Although one can only speculate given the present data, this effect might be related to a more general preference for attribute information. Factor A, input organization, exhibited no significant effect. The equality of recall times over brand and attribute inputs is compatible with the absence of differences in learning time or error rate.

The B x C and G x C interactions are probably attributable to differences in large A x C interaction for each product combined with the nature of the design. As noted earlier, any A x B interaction is confounded with G. Thus, G x C is confounded with A x B x C, and the G x C interaction may well be explained by the product differences (Factor B) in the A x C interaction that are so apparent in Figure 2. In a similar way, the B x C interaction, which is smaller, seems to have no natural interpretation. It may be related to product differences or to the large A x C effect combined with the confounding of Factors A and B within groups. In sum, the G x C and the B x C interactions show that product category affects the strength and nature of the effect of display format on memory structure. However, these differences are small when compared to the main A x C interaction, accounting for less than a fourth of its variance.

Finally, the interactions with subjects (S x A, S x C and S x A x C) should be acknowledged as statistically significant. Since these effects are not large and reflect only individual differences, they are not particularly meaningful. For completeness, it is noted that the error term is composed of the G x D and S x D interactions.

In summary, the pattern of recall times indicates a strong effect of congruence between input organization and cue type. Attribute probes led to faster recall when the input organization was attribute-based, and similarly for brand cues with brand-organized information. This result confirms the main prediction of the study. In addition, marked product differences were found when the input x probe (A x C) interaction was examined separately for each product.

Results for Individual Subjects

Chronometric studies are usually based on data analyses that are complete within individual subjects. This requires the collection of many data points to enable reliable estimates of mean response time and statistical tests with adequate power. Collecting this large amount

of data becomes more costly as the task becomes more cognitively complex. In such cases, it is not unusual to find a subject participating for 20 or more hours in one experiment (e.g., Doshier, 1976). In the present study only 16 response times are available for each subject. Thus, although statistical tests of within-subject differences can be performed, some real differences may be obscured by variability.

The 16 recall times were classified by input organization and cue type (Factors A and C), creating four groups with four observations each. For each input organization, there are two mean response times, one for attribute cues and one for brand cues. The difference between the latter two means can be used to measure congruence between memory structure and input organization. This congruence will be termed input boundness. If the product information was learned in an attribute format, the mean for attribute probes is subtracted from the mean for brand probes. The larger this difference, the greater the input boundness. These then can be thought of as estimates of the degree to which a subject's internal representation is congruent with the external source. For each subject two such measures exist, one for attribute-organized input and one for brand-organized input. By summing these two measures, individual estimates of input boundness are obtained.

Using the estimate of MSE obtained from the analysis of variance, a t-test was performed for each subject. The significance level for this t-test (one-tailed) was set at .10 to favor the detection of differences in spite of the small number of data points. Using this criterion, 8 of 20 subjects exhibited significant input boundness. Thus, the main prediction and finding of this study, the congruence between input organization and memory structure, may be characteristic of less than half the subjects.

The reduced number of subjects showing an effect of input boundness could be due, in part, to the reduced statistical power of the individual tests. Additionally, the individual estimates of input boundness are confounded with product category. The observed differences across products may weaken comparability between mean response times for a single subject. Finally, it must be remembered that typical chronometric analyses of individual subjects require much larger sample sizes than available here.

Written Free Recall

After both sets of product information had been learned and probed via the cued speeded recall task, a written free recall was collected. For each product class, an empty 4 x 4 matrix was shown. The row and column labels were provided, and subjects were required to write in the 16 missing cell values. Each such output sequence consists of 15 intercell transitions.

Each of these transitions can occur within a column, within a row, or both to a different row and a different column. Since the booklets arranged information by column, a column transition was judged to be congruent with the original display. Row transitions could only occur if the memory structure is dissimilar to the original display. Therefore, each transition was coded as congruent (column) or incongruent (row or other) with the original input organization. At the end of a column or row, when all entries are filled, no more transitions within that file are possible. Thus, transitions from filled files could not be used to measure congruity with input, and they were excluded from the analysis. This left a data base of 12 transitions for each protocol.

The overall proportion of congruent transitions is .72 (344 of 480). This value indicates a strong tendency for the written free recall to be congruent with the input organization. This confirms the major finding from the recall time data, the significant tendency for memory structure to be determined by input organization.

The proportion of congruent transitions were partitioned by the two possible independent variables, input organization and probe type. These data are reported in Table 2. The striking aspect of the pattern of proportions is their uniformity except for the brand-cooking oil cell. The other three proportions are all quite high, indicating input boundness. But when cooking oil information had been learned in a brand-organized format, considerable reorganization occurred. That is, subjects tended to reorganize from a brand-based input to an attribute-based memory structure. Again, this result confirms the findings from the recall time data.

TABLE 2
Proportion of Free Recall Transition Congruent with Input Organization

Product	Input Organization	
	Attribute	Brand
Air Conditioner	.83 ^a	.78
Cooking Oil	.80	.45

^aEach proportion is based on 120 observations.

At this point it is worth recalling the problems associated with the written free recall data that were discussed in the method section. The problematical nature of this recall technique, as it was implemented in this experiment, implies that all findings must be tempered with qualification. It is comforting that the free recall data averaged over subjects agree with the results based on the cued speeded recall task. However, a stricter test of the convergent validity between these two recall techniques can be based on the data of individual subjects. The proportion of congruent transitions was computed for each subject and correlated with the Bound statistic that is based on recall times and discussed earlier. Both are measures of input boundness. The correlation over the 20 subjects was .33, a value significantly greater than zero ($p = .02$) but not very high. It should be concluded that the free recall transitions are not as reliable as they should be--or as they could be with a different experimental procedure. However, when averaged over subjects, they probably present a satisfactorily accurate picture of memory organization.

Discussion

Summary of Results

The main finding was the congruence between input organization and memory structure. Averaged over the twenty subjects, both cued recall times and the output order of the written free recall supported this conclusion. Individual subject analysis of recall times suggested that 10 of the 20 subjects were significantly input bound. Input boundness was sensitive to type of product, with air conditioners inducing relatively consistent input boundness and cooking oil exhibiting more reorganization of the memory structure.

An important question is whether people prefer attribute-based or brand-based memory structures. A

very preliminary result of the present study is that attribute structures were preferred. This result is as intriguing as it is tentative.

Generality

The findings of this study, although clearly supporting the view that memory structure is input bound, should be considered only in light of product differences. There was much greater reorganization (less input boundness) for cooking oil than for air conditioners. Possible factors that can influence the input boundness of a product class need to be isolated. It is interesting to note, for example, that both the brand names and the attribute names were more familiar for cooking oil. Also, at least one of its attributes (polyunsaturated) fat count is health-related and apt to be especially salient. Clearly, a wider range of products will have to be examined before any relation between reorganization potential and product type can be established.

Because the subjects in this experiment were undergraduates, the findings may not generalize to representative shoppers. Bettman and Kakkar (1977) argue that such shoppers would be more brand conscious than undergraduates. A shopper's experience within a particular product category may contribute to brand consciousness or, alternatively, to the ability to reorganize by attributes. Questions of this nature can only be answered by studies designed to test the generality of the present findings over a much wider sample of products and subjects.

The Chronometric Technique

One of the positive results of this study is the demonstration of the usefulness of chronometric analysis. This is especially true for the study of the representation of product information in memory. The present use of chronometric analysis to identify the memory structure of product information parallels the recent recognition of the importance of this topic (e.g., Olson, 1977 and Bettman, 1978). However, chronometric analysis can be applied to a wide range of other consumer research problems. See Gardner, Mitchell and Russo (1978) for a discussion of this point and an application of this technique to low involvement advertising. As a final methodological note, the reader should be reminded that the chronometric technique used here is among the least sophisticated. More powerful methods are available and, hopefully, will be brought into the study of consumer behavior. However, even the simple chronometric techniques may be adequate given the size of the effects observed here which are considerably larger than those reported in most chronometric studies.

Implications for Policy

Memory structure for product information, and its relation to input format on the one hand and decision rules on the other, has important policy implications. We give one example. To the extent consumers are input bound, their memory structures will reflect the organization of product information in the environment. As just pointed out, this organization is largely brand-based (advertising, etc.). This is all well and good if consumers make better purchase decisions using brand-based strategies. But if attribute-based strategies are superior, and a growing body of evidence suggests this (e.g., Russo and Doshier, 1976), then brand-organized storage contributes to less accurate and more effortful purchase decisions. We may then find a conflict between superior decision rules and the organization of currently available product information.

This suggests a potential need for public policy intervention. More importantly, it suggests that any policy of information provision may want to alter the current format of product information. If the ultimate policy goal is to improve consumer decisions through information disclosure, the format of the information should be designed for compatibility with the preferred decision rules.

Future Research

A major direction for future research is to examine this relation between memory structures and decision rules. Does the structure of stored product information determine the decision rule(s) that can be employed? Or, conversely, does a preferred decision rule lead to the reorganization of product information into a compatible memory structure? The relative strength of influence between memory structure and decision rules is an important question in consumer research.

A related issue concerns the conditions surrounding the acquisition of knowledge about products. Consumers may intentionally retain information with the conscious goal of making a purchase decision. Knowledge can also be retained, possibly incidentally, for more general, future use (Greeno, 1976).

A second topic for future research is the influence of different external organizations of product information on memory structure. What is the preferred memory organization if a complete brand x attribute array is available? Suppose the externally available information is very unstructured, possibly in a discursive format or only accessible in a random order. Will shoppers attempt to organize an unstructured external display into a consistently preferred internal structure? Finally, what if product information is available as it typically is in the market place: from several sources (advertising, use experience, word of mouth, published ratings, etc.) that are not naturally comparable and that may include inconsistencies? Further, most real-world product information is brand-centered. Does this imply that memory structures are usually brand-organized whether or not attribute-based structures are preferred?

In the realworld, memory of product information is the link between information acquisition and the purchase decision. This topic is a major understudied area of consumer behavior. There are many important research questions associated with it, only a few of which have been posed above. By actively studying memory of product information, it may be possible to close a major gap between current laboratory research and the realworld consumer setting.

References

- John R. Anderson, Language, Memory and Thought (Hillsdale, NJ: Lawrence Erlbaum, 1976).
- James R. Bettman, An Information Processing Theory of Consumer Choice (Reading, MA: Addison-Wesley, 1978, forthcoming).
- _____ and Pradeep Kakkar, "Effects of Information Presentation Format on Consumer Information Acquisition Strategies," Journal of Consumer Research, 3(March, 1977), 233-240.
- William G. Chase, "Elementary Information Processes," in W. K. Estes, ed., Handbook of Learning and Cognitive Processes, Vol. 5 (Hillsdale, NJ: Lawrence Erlbaum Associates, in press).

Barbara Anne Doshier, "The Retrieval of Sentences from Memory: A Speed-Accuracy Study," Cognitive Psychology, 8(July, 1976), 291-310.

James G. Greeno, "Indefinite Goals and Well-Structured Problems," Psychological Review, 83(November, 1976), 479-491

Meryl Gardner, Andrew A. Mitchell and J. Edward Russo, "Chronometric Analysis: An Introduction and an Application to Low Involvement Perception of Advertisement," in H. Keith Hunt (ed.), Advances in Consumer Research, Vol. 5, Association for Consumer Research, 1978.

Roger E. Kirk, Experimental Design: Procedures for the Behavioral Sciences (Belmont, CA: Brooks/Cole, 1968).

Donald A. Norman, Memory and Attention: An Introduction to Human Information Processing, Second Edition (New York: Wiley, 1976).

Jerry C. Olson, "Theories of Information Encoding and Storage: Implications for Marketing and Public Policy," presented at the Conference on the Effect of Information on Consumer and Market Behavior (Carnegie-Mellon University, May, 1977).

John W. Payne, "Heuristic Search Processes in Decision Making," in Beverlee B. Anderson, Ed., Advances in Consumer Research, Vol. 3, Association for Consumer Research, 1976.

Herbert A. Simon and John R. Hayes, "The Understanding Process: Problem Isomorphs," Cognitive Psychology, 8(April, 1976), 165-190.

_____ and Alan Newell, "Thinking Processes," in David H. Krantz, R. Duncan Luce, Richard C. Atkinson, and Patrick Suppes, Eds., Contemporary Developments in Mathematical Psychology, Vol. 1, (San Francisco: Freeman, 1974).

Saul Sternberg, "Memory Scanning: Mental Processes Revealed by Reaction Time Experiments," American Scientist, 57(1969), 421-457.

Endel Tulving and Gordon H. Bower, "The Logic of Memory Representations," in Gordon H. Bower, Ed., The Psychology of Learning and Motivation, Vol. 8 (New York: Academic Press, 1974).

LOW INVOLVEMENT VERSUS HIGH INVOLVEMENT COGNITIVE STRUCTURES¹

John L. Lastovicka, Temple University
David M. Gardner, University of Illinois

Abstract

Multidimensional scaling of compact car perceptions is used to examine differences in cognitive structure between those highly involved with compact cars and those who are not. The empirical results, in line with prior theory, suggest a less differentiated and integrative structure for the low involved.

Introduction

Krugman's (1965) low involvement learning model has recently received renewed attention by consumer behavior researchers (Maloney, 1977; Banks and Hart, 1977). Krugman contended that television advertising is a special low involvement communication situation in which receiver responses are akin to the passive learning of nonsense syllables. He further suggested that the repetition of advertising resulted in a replacement of old brand perceptions with a new set of beliefs. This new cognitive structure was said to guide brand choice behavior without changing attitude (affect) first.

Krugman's original research stimulated consumer behavior researchers to conceptualize the advertising process in a situation-specific "micro theoretical" manner. For example, Ray et al. (1973) present alternative hierarchies of effect for different levels of involvement. One hierarchy, the standard COGNITIVE→AFFECTIVE→CONATIVE order, is labeled the learning hierarchy, and is seen as most appropriate for high involvement decisions which typically deal with high priced, high risk products. Another hierarchy, a COGNITIVE→CONATIVE→AFFECTIVE order, is labeled the low involvement hierarchy, and is seen as being appropriate for much repetitive brand choice behavior of inexpensive, low risk products. Figure 1 presents the two hierarchies.

Figure 1
Alternative Hierarchies of Effect[†]



[†]Adapted from Ray et al. (1973). "Learn-Feel-Do" terminology from Raymond (1976).

Despite the extensive research on low involvement consumer behavior (Maloney, 1977; Banks and Hart, 1977; Bowen and Chafee, 1974; Chafee and McLeod, 1973; Hupfer and Gardner, 1971; Preston, 1970; Robertson, 1976; Rothschild and Ray, 1974; and Ward, 1975), very little is known about the details of the differences between the cognitive structures in the two alternative hierarchies of effects.

One general assumption in this literature is that a low involvement cognitive structure is much less complex than a high involvement cognitive structure. Based on this assumption, it has been suggested that those advertising low involvement products discuss fewer product attributes in their advertising copy (Rothschild, 1977; Lastovicka and Gardner, 1977). Yet this global hypothesis, of less complex cognitive structure in low involvement, has not been specifically put to empirical test. The intent of this paper is to test this global hypothesis.

Method and Hypotheses

One approach of demonstrating the effects of involvement on cognitive structure is to compare and contrast groups of respondents who differ in their level of involvement with a given product class. In this study subjects highly involved with compact cars were compared to subjects who had a low level of involvement with compact cars. Although a correlational approach necessarily creates ambiguous statements, such an approach provides an opportunity to richly demonstrate the powerful effects of involvement on cognitive complexity.

Schroder, Driver and Streufert (1967) discuss several aspects of cognitive complexity in information processing. Two important aspects are "differentiation" and "integration." Broadly defined, differentiation refers to dimensionality or the number of dimensions used by an individual in processing information; and integration refers to the extent dimensions are interrelated or used simultaneously. The influence of involvement on both differentiation and integration were tested with these hypotheses:

H₁: A unidimensional model of car model similarities will produce a better representation of car model perceptions for low involved consumers than for high involved consumers. More complex multi-dimensional representation of brand perceptions will be better models of the cognitive structure of the high involved, but not necessarily the low involved consumers.

H₂: For a given dimensionality, highly involved consumers will tend to be more integrative and rely simultaneously on several perceptual dimensions. Low involved consumers will rely primarily on a lesser number of perceptual dimensions.

¹Preparation of this paper was aided by University of Illinois participation in the General Motors Intercollegiate Marketing Competition. The authors acknowledge with thanks the assistance of graduate students Patrick Merrill, Bruce Newman, Gary Pope, and Kwok-Cheung Wong for use of their data.

Data

One hundred twenty-seven respondents from an introductory marketing course completed a questionnaire dealing with low price, medium price and luxury domestic compact cars. Responses were collected on:

1. Twenty-two lifestyle items. A 1-5 strongly disagree --strongly agree scale was used.
2. Forty-five paired comparison dissimilarity items for ten different compact cars. A 1-9 extremely similar --extremely dissimilar scale was used to measure perceptions of these ten compact car models:

Pontiac Ventura	AMC Pacer
Ford Maverick	AMC Hornet
Mercury Comet	Oldsmobile Omega
Buick Skylark	Ford Granada
Mercury Monarch	Chevrolet Nova

These ten cars include the "look-alike" parity models of Granada-Monarch, Nova-Omega-Skylark-Ventura, and Maverick-Comet.
3. In addition to the forty-five paired comparisons for the ten stimuli, a repeated set of ten paired comparison items were completed by each subject. Test-retest reliability on the paired comparison task could then be measured for each subject.
4. A battery of familiarity questions. Subjects were asked how familiar they were with each of the ten car models on a 1-7 not at all familiar--very familiar scale.
5. A battery of demographic questions including age, sex and perceived social class.
6. A battery of ten questions in which respondents were asked to sort each of the car models into different categories of acceptability. The question posed was:

"How acceptable to you are each of the following car models?"

	Definitely Not Acceptable			Neutral		Definitely Acceptable	
	1	2	3	4	5	6	7
Pontiac Ventura	1	2	3	4	5	6	7
AMC Hornet	1	2	3	4	5	6	7
.
.
.
Chevy Nova	1	2	3	4	5	6	7

This last battery of ten questions is based on the Sherif and Sherif (1967) Own Categories Procedure measurement of involvement. The Sherif method starts with a list of items relevant to an attitude. This may be objects, pictures, or verbal statements about some topic or issue. In the current research the items are ten brand names of compact cars.

In the Own Categories Procedure respondents are asked to sort items into categories of acceptability. Considerable research (Sherif and Howard, 1961; Sherif, Sherif and Nebergall, 1965) confirms that the items that an individual accepts, rejects and towards which he is neutral or noncommittal varies systematically with his personal involvement. Specifically the Sherifs found:

Proportional to his lack of involvement, the number of positions the individual accepts and rejects become approximately equal and his latitude of noncommitment increases. This means that highly involved persons have a much broader latitude of rejection than persons less concerned, and that they remain noncommittal toward fewer

positions, even when not required to evaluate all of them. (Sherif and Sherif, 1967, p. 191, italics added.)

Typically, an involvement index is created in which the ratio of the number of items rejected to the number of items accepted, adjusted by the number of neutral items, is computed. With the current data such a measure was undefinable for many respondents for whom the number of items accepted and/or rejected was zero.

However, an involvement index which relied on the number of items which subjects were neutral or noncommittal about was operable with the current data. Depending on their degree of noncommitment or neutrality, respondents were rated in involvement with the following index:

$$\text{Involvement Index for Individual}_i = \frac{10}{\sum_{j=1}^{10} |X_{ij} - 4|}, \quad (1)$$

where X_{ij} is the i th individual's response to the j th of the ten acceptability questions and 4 is the neutral point on the 1-7 scale. Following Sherif, a high involvement index number computed from (1) indicates a small degree of noncommittal and a high level of involvement. A low index number represents a large latitude of non-committal and a corresponding low level of involvement.

By recoding the responses to the acceptability questions for the ten makes and models of compact cars using (1), an involvement index with a range of 0-30 was computed for each respondent.

Analysis and Results

A pair of multidimensional scaling (MDS) models, Young's TORSCA (1968) and Carroll and Chang's INDSCAL (1970), were used to uncover the underlying dimensions used by the respondents in evaluating the similarities between the ten compact cars.

Information processing researchers Schroder, Driver and Streufert (1967) cite aggregate level MDS methods as a good measure of differentiation or the number of dimensions used by individuals in processing information. Individual differences MDS models such as INDSCAL, which indicate to what degree respondents use the underlying dimensions, offers direct measurement of integration or the degree to which the dimensions are used simultaneously and are interrelated. Also, MDS researchers Shepard, Romney and Nerlove (1972) show MDS useful in recovering degree of cognitive complexity. MDS, therefore, seems well suited for tapping cognitive complexity.

Preprocessing the Data

Since each respondent provided two sets of judgments for ten of the possible forty-five paired comparisons, a reliability index was computed for each respondent with Spearman's rank order correlation. Those 79 respondents of the 127 whose reliabilities were above the critical level of .54 were retained for both the TORSCA and INDSCAL analyses. The average Spearman correlation for this retained group of 79 was .77.

Once involvement indices were computed, the retained sample of 79 was divided into three, approximately equal sized, involvement groups: low, medium, and high. The low group consisted of thirty respondents whose involvement indices were 17 or less. Twenty respondents, with indices between 18 and 21, were the medium group; and another twenty-nine respondents, with indices of 22 and over, were the high group. The mean involvement indices

for the low, medium and high groups were 14.28, 19.90 and 24.70, respectively.

Hypothesis One

Average, standardized similarity matrices for the paired comparison data were computed for the high and low involvement groups. The result was two 10 x 10 average matrices of car model dissimilarities. The high and low involvement dissimilarity matrices were scaled at an aggregate level using the TORSCA algorithm. In each case, solutions were obtained in one through four dimensions. Figure 2 shows the relationship between the Kruskal's stress statistic (the goodness of fit between the original dissimilarities data and the n-dimensional MDS configuration) and the dimensionality of the solution. Examination of Figure 2 should be done in the same spirit as the "root staring" procedure in factor analysis in which eigenvalues are plotted versus their number. This comment is made because established statistical tests for comparing stress levels are not available. Just as in "root staring," then, subjective judgment must play a large role in the current analysis.

The stress measures for the high and low involvement scalings shown in Figure 2 are in support of the first hypothesis. A simple unidimensional model of dissimilarities is a much better representation of cognitive structures for the low involvement group than the high involvement group. This simplest scaling indicates that a more complex model is needed for the high involvement group. For the two and three dimensional scaling solutions, differences in stress are not great. Yet these two and three dimensional solutions offer slightly better fit for the high involvement group. Finally, for the four dimensional scalings, the most complex models built in this analysis, the results are as theory predicts. Such further complication does not provide a better representation of cognitive structure for the low involvement group, yet better fit is marginally obtained for the high involvement group.

Hypothesis Two

The INDESCAL model was used to examine cognitive complexity in terms of integration. This model assumes a common stimulus space, with differential weighting of the dimensions of this common space for each respondent. A respondent's position in the INDESCAL person space represents the salience he assigns to each of the dimensions in the common stimulus space. Thus, the weights can be used to estimate an individual stimulus space for each person. The individual stimulus configurations are based on the common stimulus space, but are differentially "stretched" in accord with the square roots of the respondent's own weights.

The two dimensional group stimulus space in Figure 3 was interpreted with the classic LUXURY and SPORTY dimension.

The person space, showing each of the 79 respondent's saliences for the two dimensions of the common stimulus space, is shown in Figure 4. Using visual clustering the respondents were grouped into three clusters representing three integrative styles. A "left" cluster of 29 respondents who made disproportionate use of the SPORTINESS dimension was identified. A "center" cluster of integrative respondents who made roughly equal use of both of the common stimulus space dimensions has 24 respondents. The third cluster, a "right" cluster of 26 respondents, contains those who made disproportionate use of the LUXURY dimension.

The next step of the analysis was to test the influence of involvement on integrative complexity. In a two group discriminant analysis, involvement along with life-

style, familiarity and demographic measures were used to predict membership in the integrative "center" cluster or the nonintegrative "left" and "right" cluster. Though the prime interest is in detecting the influence of involvement on integration, it was felt the predictive power of this variable should be compared with others. Examination of normalized discriminant function weights allows easy comparative assessment of the predictor variables.

Rather than use the entire battery of the 32 lifestyle and familiarity measures as predictors in the discriminant analyses, two independent principal axes factor analyses were used on these measures to create a set of six parsimonious underlying factors.

The first factor analysis conducted on the 22 lifestyle items revealed four common factors. After a varimax rotation the four factors were easily interpretable. The first factor, an automotive knowledge factor, had high positive loadings on items such as: "I know a lot about cars;" and "People often come to me for information." The second factor, a functional automotive preference factor, had high positive loadings on items like: "I like to drive a car that will hold up in an accident;" and high negative loadings on items such as: "You can tell a lot about a person from the model of car he drives." The third factor, an image preference factor, loaded highly on items like: "I think cars are a mark of status;" and "I like to have the best looking car on the road." The fourth lifestyle factor, an automotive thrill-seeking factor, loaded highly on items such as: "I like to drive fast;" and "I like to listen to music while driving." Factor scores were then estimated on these four lifestyle factors.

The second factor analysis was conducted on the familiarity ratings for the ten compact cars. Two factors were extracted in this analysis; the first was an overall familiarity factor and the second was familiarity with certain models of parity cars. Familiarity with the differences between the "look-alike" parity models of Omega, Nova, Skylark and Ventura, for example, was independent of overall familiarity. Factor scores were also estimated for these two familiarity factors.

Using lifestyle factor score estimates, familiarity factor score estimates, the involvement index, age, social class and a dummy variable for sex, a significant discriminant function was found to differentiate between the integrative and non-integrative groups. The obtained value of Rao's F-ratio approximation, $F_{10,68} = 1.96$, is just significant at the .05 level. Using the discriminant function weights shown in Table 1, 70.8 percent of the integrative group and 72.7 percent of the non-integrative group could be correctly classified. Predictive validity was further tested using U-method pseudo jack-knife classification (Crask and Darden, 1977). With the U-method a classification function was computed for each of the 79 cases with that case omitted from the computations. Since each of these functions was used to classify the left-out case, a less biased classification occurs.

Based on the U-method jack-knifed classification functions, 50 percent of the integrative group and 67.3 percent of the non-integrative group were correctly classified.

Examination of the relative size of each standardized discriminant weight in Table 1 gives an indication of the relative importance of the ten variables in differentiating between the two groups. As hypothesized, involvement is found to be a good explanatory variable for cognitive integration. The weight for the involvement index, .806, is almost twice the size of the next largest weight. Examination of the cluster means on the discriminant function shows that the "center" cluster,

Figure 2
A Comparison of the Kruscal Stress
Statistic Across 1-4 Dimensional
Solutions for Group TORSCA Scalings

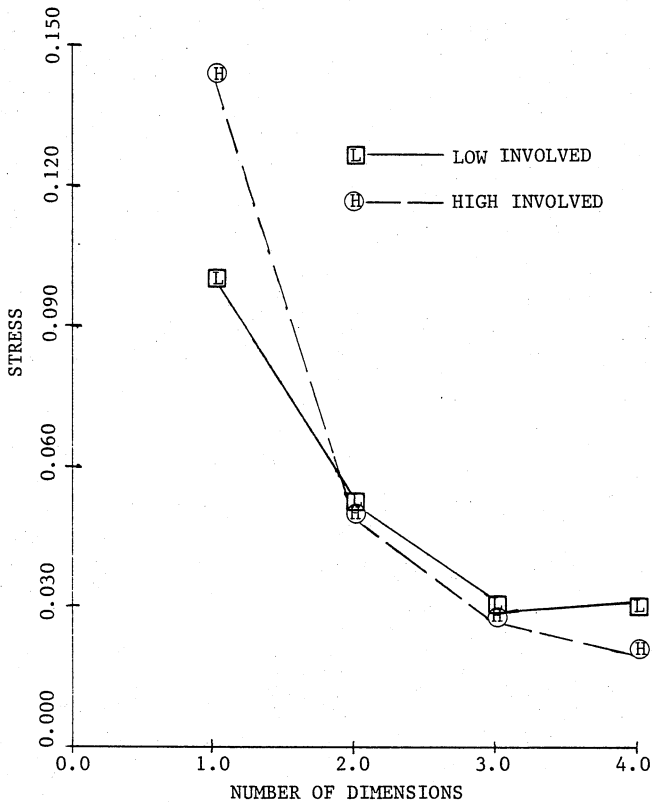


Figure 3
Two-Space Common Stimulus Con-
figuration for INDSCAL Scaling Analysis

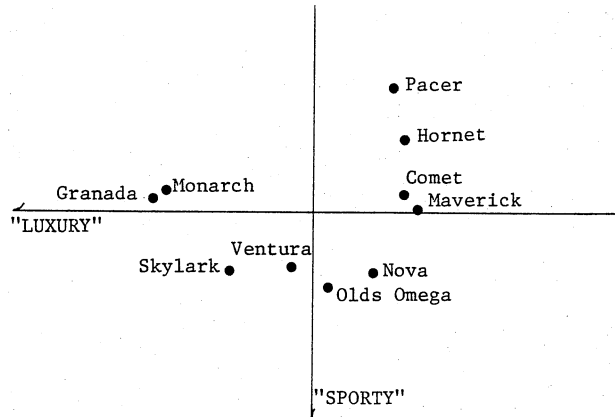


FIGURE 4

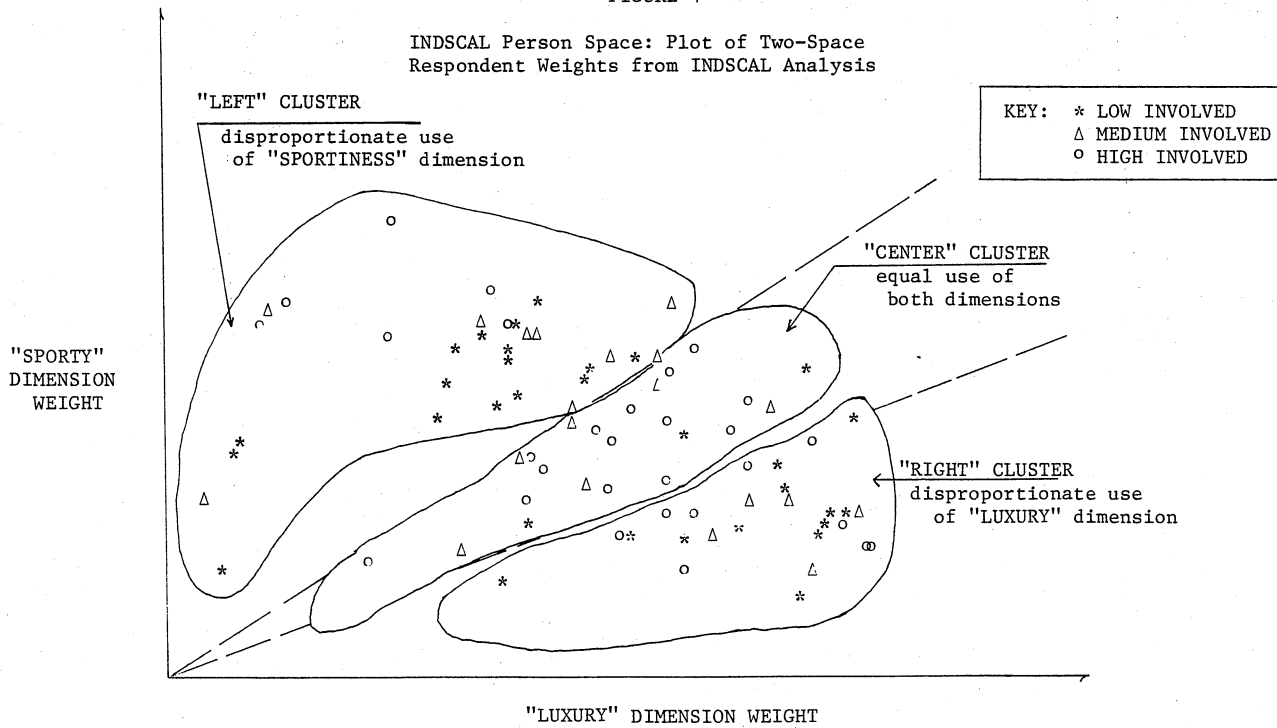


Table 1
Standardized Discriminant Weights

Predictor Variable	Weight
Automotive Knowledge	.045
Functional Automotive Preference	-.493
Automotive Image Preference	-.181
Automotive Thrill Seeking	.244
Overall Compact Car Familiarity	-.099
Parity Compact Car Familiarity	-.317
Involvement Index	.806
Age	-.154
Sex	.055
Social Class	.031

Group Means on Discriminant Function

Group	Mean
Integrative ("Center" Cluster)	2.99
Non-Integrative ("Left" and "Right" Clusters)	1.84

the integrative group which used both the INDSCAL stimulus space dimensions equally, has the highest score. It seems then that equal use of both dimensions occur primarily under high involvement. Low involvement cognitive structures seem less integrative and rely primarily on one dimension.

Discussion

The general findings of this study are in line with prior theory. Low involvement cognitive structures do seem to be simpler than high involvement structures in at least two ways. First, low involvement structures seem less differentiated as they can be represented adequately with fewer dimensions than high involvement structures. Second, low involvement structures tend to be less integrative. In the current data, a two space map of a low involved individual's compact car perceptions is typically most reliant on one dimension. The simultaneous, integrative approach is apparently not worth the effort on the part of the low involved consumer.

Despite the support for the global hypothesis, several points must be kept in mind.

First, the differences found in cognitive structure can only be said to be potentially due to involvement. Crucial differences besides involvement may be responsible for the observed phenomena. Further research in this area should include both use of experimentation and multiple measurement approaches to involvement.

Second, cognitive structure, the dependent measure, was measured only along two dimensions: differentiation and integration. Cognitive differences in terms of discrimination, for example, were not examined.

Third, differences have been examined between individuals for a given product. Though such an approach is useful for market segmentation, the real thrust of low involvement consumer behavior concerns differences between products.

In conclusion, this study should be seen as a very basic exploration into the nature of the differences between high and low involvement cognitive structures. The paper should help to underscore that there is a difference between the low involved consumer and the high involved consumer that most researchers have implicitly assumed.

References

Banks, S. and E. W. Hart, "Advertising and Promotional Methods," in Robert Ferber, ed., A Synthesis of Selected Aspects of Consumer Behavior (Washington, D.C.: National Science Foundation, 1977, in press).

Bowen, L. and S. H. Chafee, "Product Involvement and Pertinent Advertising Appeals," Journalism Quarterly, 51 (Winter, 1974), 613-21.

Caroll, J. D. and J. J. Chang, "Analysis of Individual Differences in Multidimensional Scaling Via an N-Way Generalization of Eckart-Young Decomposition," Psychometrika, 35 (1970), 283-319.

Chafee, S. H. and J. M. McLeod, "Consumer Decisions and Information Use," in S. Ward and T. S. Robertson, eds., Consumer Behavior: Theoretical Sources (New York: Prentice-Hall, 1973).

Crask, M. R. and W. D. Perreault, Jr., "Validation of Discriminant Analysis in Marketing Research," Journal of Marketing, 14 (February, 1977), 60-8.

Hupfer, N. and D. Gardner, "Differential Involvement With Products and Issues: An Exploratory Study," Proceedings of the Association for Consumer Research, Second Conference, 1971.

Krugman, H. E., "The Impact of Television Advertising: Learning Without Involvement," Public Opinion Quarterly, 29 (Fall, 1965), 349-56.

Lastovicka, J. L. and D. M. Gardner, "Components of Involvement," in John C. Maloney, ed., Attitude Research Plays for High Stakes (Chicago: American Marketing Association, 1977, in press).

Maloney, John C., ed. Attitude Research Plays for High Stakes (Chicago: American Marketing Association, 1977, in press). (Proceedings of the American Marketing Association's Eighth Annual Attitude Research Conference. All papers from two sessions: "Must Consumers Think Before They Act?--A Review of Low Involvement Theory" and "Low Involvement Versus High Involvement Situations--Some Guidelines for Action," deal with the involvement issue.)

Preston, L. L., "A Reinterpretation of the Meaning of Involvement in Krugman's Models of Advertising Communication," Journalism Quarterly, 47 (Summer, 1970), 287-95.

Ray, M. L., A. G. Sawyer, M. L. Rothschild, R. M. Heeler, E. C. Strong, and J. B. Reed, "Marketing Communication and the Hierarchy of Effects," in P. Clarke, ed., New Models for Mass Communication Research (Beverly Hills, Cal.: Sage, 1973).

Raymond, C. Advertising Research: The State of the Art (New York: Association of National Advertisers, 1976).

Robertson, Thomas S., "Low Commitment Consumer Behavior," Journal of Advertising Research, 16 (April, 1976), 19-27.

Rothschild, M. L., "Advertising Strategies for High Involvement and Low Involvement Situations," in John C. Maloney, ed., Attitude Research Plays for High Stakes (Chicago: American Marketing Association, 1977, in press).

Rothschild, M. L., and M. L. Ray. "Involvement and Political Advertising Effect: An Exploratory Experiment," Communication Research, 1 (July, 1974), 264-84.

Schroder, H. M., M. J. Driver and S. Streufert, Human Information Processing (New York: Holt, Rinehart and Winston, 1967).

Shepard, R. N., A. K. Romney and S. B. Nerlove, eds., Multidimensional Scaling, Volume I. (New York: Seminar Press, 1972).

Sherif, M. and C. I. Hovland. Social Judgment: Assimilation and Contrast Effects in Communication and Attitude Change (New Haven: Yale University Press, 1953).

Sherif, M. and C. W. Sherif, "The Own Categories Procedure in Attitude Research," in Martin Fishbein, ed., Readings in Attitude Theory and Measurement (New York: Wiley, 1967).

Sherif, W. C., M. Sherif and R. E. Nebergall. Attitude and Attitude Change: The Social Judgment Involvement Approach (New Haven: Yale University Press, 1965).

Ward, Scott, "Marketing Strategies for 'Low Ego-Intensity/Low Risk' Products," Intercollegiate Case Clearing House, 1975.

Young, F. W., "TORSCA, An IBM Program for Nonmetric Multidimensional Scaling," Journal of Marketing Research, 5 (August, 1968), 319-21.

A STUDY OF THE INTERFACE BETWEEN ATTITUDE STRUCTURE
AND INFORMATION ACQUISITION
USING A QUESTIONNAIRE-BASED INFORMATION-DISPLAY SHEET

Morris B. Holbrook, Columbia University¹
Karl A. Maier, Ketchum, MacLeod & Grove, Inc.

Abstract

A new technique for measuring information acquisition in a survey questionnaire is applied to a study of the relation between the selection of brand-attribute cues and the multiattribute model of attitude structure. A relationship is found between attribute importance and the acquisition and impact of product-related information.

Introduction

A growing body of consumer research has investigated the acquisition of information during the decision-making process. By far the most common technique for studying this search phenomenon uses information-display boards that array cues on m attributes of n brands in an $m \times n$ matrix (Bettman, 1976; Bettman and Jacoby, 1976; Chestnut, 1976; Jacoby, 1975, 1976; Jacoby, Chestnut, and Fisher, n.d.; Jacoby, Chestnut, Weigl, and Fisher, 1976; Jacoby, Szybillo, and Busato-Schach, 1977; Payne, 1976a, b). Typically, choice-simulating subjects select a brand-attribute combination of interest, remove a card from that cell of the matrix, read the information on the back of the card, and deposit it into a pile before choosing the next card. The resulting deck of used cards represents the extent and sequence of information sought about each brand on each attribute.

The potential relevance of such information-acquisition measures to the multiattribute models of attitude structure is obvious. Yet a comprehensive review of recent attitude research in marketing suggests that little effort has been made by consumer researchers (with or without the use of information-display boards) to study the interface between attitude structure and the informational inputs that presumably shape beliefs or other basic components of the multiattribute models (Holbrook, 1976).

Holbrook (1975, 1976, 1977) has argued that--according to Bayesian decision theory (Edwards, 1965; Green, 1964; Winter, 1975), Berlyne's (1960) theory of cognitive motivation, and the theory of perceived risk (Cox, 1967; Ross, 1975; Taylor, 1974)--information to aid consumer decision making should be sought primarily on those attributes that are subjectively perceived as most important. More broadly, it appears that the value of information should depend upon the importance, uncertainty, and confidence associated with the attributes to which it pertains. This conclusion is supported by the so-called Principle of Information-Processing Parsimony (Haines, 1974, p. 96), which suggests that limited cognitive capacity forces problem solvers to adopt heuristics restricting their attention to the minimum amount of data necessary for satisficing decisions (Bettman, 1974; Haines, 1974; Payne, 1976a; Wright, 1973, 1974). That principle is consistent with the lexicographic decision model's assumption that information on various attributes will

be sought in order of their perceived importance (Payne, 1976a, b; Wright, 1975), but apparently this hypothesized relation between subjective attribute importance and search priority has been tested only occasionally, rather unsystematically, and with mixed results (Jacoby et al., n.d., 1976; Bettman, 1974; Myers, 1976; Nakanishi and Bettman, 1974; Tigert, 1966; Winter, 1975). In further accord with the Principle of Parsimony, numerous studies have shown that a very simple model of attitude structure, summing only the belief scores on those attributes considered the few most important, predicts affect at least as well as more complex forms of the multiattribute model (Bass and Wilkie, 1973; Holbrook, 1976; Holbrook and Hulbert, 1975). All these perspectives on attribute importance suggest that its role at the interface between attitude structure and information acquisition may be fundamental to our understanding of preference formation and brand choice.

If such propositions are to be tested on reasonably large and representative consumer samples of the type recently called for by Ferber (1977), it will be useful to develop measures of information acquisition that can be mass administered through survey-questionnaire techniques. To this end, the authors have designed an information-display sheet that is analogous to the information-display board except that the respondent acquires cues by removing gummed stickers placed over the cells of interest in an $n \times m$ brand-attribute matrix. Self-administration of this instrument makes it suitable for use in field-survey research.

Accordingly, the study reported below had three major purposes: (1) to evaluate the feasibility of the questionnaire-based sticker-removal task as a measure of information acquisition; (2) to test the hypothesis that attribute importance, as an indicant of information value, is a key determinant of the extent and sequence of information seeking on that attribute; (3) to test the hypothesis that preference is adequately predicted by a simple model summing the evaluative beliefs on those attributes considered important enough to be the object of search.

Method

Subjects and Choice Task

One hundred students at Columbia University's Graduate School of Business participated in a questionnaire-based choice task in which each respondent sought information before selecting a phonograph record to receive as a raffle prize if he was one of five winners in the lottery for a \$5.00 gift certificate. In order to approximate the conditions typical of a mail survey, respondents completed the questionnaires at their leisure on a self-administered basis.

The records presented in the choice task were fictitious vocal albums that differed on six attributes described by the questionnaire as follows:

1. SINGER'S STYLE: traditional--a well known pop artist who has been around for a number of years; or contemporary--a relative newcomer setting trends in the current idiom

¹ Morris B. Holbrook gratefully acknowledges the support of Columbia University's Faculty Research Fund.

2. TYPE OF SONGS: standards--familiar tunes that most listeners have heard before; or originals and recent hits--new songs composed by the performer or recently made popular by someone else
3. TYPE OF PRODUCTION: studio--recorded in a fancy studio, usually with complex background arrangements; or live--recorded in front of a live audience, typically with a small back-up band
4. JACKET: informative--album cover contains extensive liner notes with biographical data on the artist and full credits for producers, arrangers, musicians, etc.; or visual--album cover features attractive photos or other graphic artwork suggesting the mood conveyed by the recording inside
5. LABEL: major--one of the big entertainment-industry conglomerates; or independent--a small, privately run record company
6. PRICE: \$5.59--a 20% discount from the list price of \$6.98; or \$4.19--a 40% discount from the list price of \$6.98

Attribute-Specific Measures

Immediately following these descriptions, the six attributes were rated on 7-point scales representing: (1) the importance of each attribute to the respondent's evaluation of a pop vocal album (from "not at all important" to "extremely important"); (2) the respondent's confidence in his ability to understand and interpret information concerning each attribute (from "not at all confident" to "extremely confident"); (3) the respondent's certainty that virtually any vocal recording on the market would be satisfactory with respect to each attribute (from "not at all certain of satisfaction" to "extremely certain of satisfaction"). In addition, the two characteristics cited in describing each attribute were rated for their desirability on the 7-point scales shown below:

	EXTREMELY UNDESIRABLE	EXTREMELY DESIRABLE
Traditional singer	_ : _ : _ : _ : _ : _	_ : _ : _ : _ : _ : _
Contemporary singer	_ : _ : _ : _ : _ : _	_ : _ : _ : _ : _ : _
Standard songs	_ : _ : _ : _ : _ : _	_ : _ : _ : _ : _ : _
Originals and recent hits	_ : _ : _ : _ : _ : _	_ : _ : _ : _ : _ : _
Studio recording	_ : _ : _ : _ : _ : _	_ : _ : _ : _ : _ : _
Live recording	_ : _ : _ : _ : _ : _	_ : _ : _ : _ : _ : _
Informative cover	_ : _ : _ : _ : _ : _	_ : _ : _ : _ : _ : _
Visual cover	_ : _ : _ : _ : _ : _	_ : _ : _ : _ : _ : _
Major label	_ : _ : _ : _ : _ : _	_ : _ : _ : _ : _ : _
Independent label	_ : _ : _ : _ : _ : _	_ : _ : _ : _ : _ : _
Price of \$5.59	_ : _ : _ : _ : _ : _	_ : _ : _ : _ : _ : _
Price of \$4.19	_ : _ : _ : _ : _ : _	_ : _ : _ : _ : _ : _

The importance and confidence ratings were scored from 1 to 7. By contrast, the certainty ratings were scored in the direction of increasing uncertainty, from 7 to 1. Finally, in acknowledgement of the debate surrounding the proper coding of evaluative scores (Ahtola, 1975; Fishbein, 1976; Glassman and Fitzhenry, 1976; Lutz, 1976), two alternative codings of the desirability ratings were investigated: -3 to +3 and 1 to 7.

Test Objects

The dichotomous characteristics for each attribute were used to specify eight records according to the kind of fractional factorial design described by Green (1974) as suitable for conjoint-measurement studies. The rows of the resulting 8x6 matrix were randomized and labelled from "A" to "H." The entire matrix of attribute-specific cues describing each record album appears in Table 1.

The Information-Display Sheet

After providing the attribute-specific ratings discussed above, the respondent turned to an information-display sheet consisting of the matrix shown in Table 1 with each cell covered by a gummed sticker appropriately labelled from A1 to H6. Written instructions asked the respondent to remove enough stickers to give him the information needed to choose the album he wished to receive as his raffle prize if he won the drawing for a \$5.00 gift certificate. The bottom of the page contained another matrix entitled "Order of Sticker Removal" with boxes labelled from "1st" to "48th." The respondent was instructed to place each sticker into the box designating its order of removal. He was told that he could obtain all the information needed to make his choice by peeling off as many stickers as he wanted, but that each piece of information acquired would result in the deduction of 5 cents from the \$5.00 value of his potential prize.

After completing this task and making his choice, the respondent was further instructed to continue removing stickers until he had collected enough information to permit him to rank all eight recordings in order of preference. He was told that no cost would be imposed for these additional stickers and that they should therefore be placed in a second set of boxes labelled "1st" to "48th" and entitled "Order of Removal of Additional Stickers." When he had completed this second process of cue selection, the respondent ranked all eight records in order of preference.

In sum, this procedure obtained measures of both choice and preference order with associated indices of costly and costless information acquisition. Not surprisingly, the 5-cent charge for cues in the choice task was somewhat discouraging to information acquisition. On the average, respondents acquired only 9.02 cues in this task, as compared with an average of 25.46 in both the choice and ranking tasks combined. Full analyses of both the single and combined information-acquisition measures produced virtually identical results. Accordingly, only the data for the combined total of cues selected are discussed in the present report.

Operational Definitions

Extent of information acquisition. The extent of information acquisition (EIA) on each of the six attributes was defined operationally as the number of cues acquired across all eight records:

$$EIA_j = \sum_{i=1}^8 ACQ_{ij} \quad (1)$$

where ACQ_{ij} is a zero-one dummy variable representing the removal of a sticker in either the choice or ranking task, i refers to the eight records ($i = 1, \dots, 8$), and j refers to the six attributes ($j = 1, \dots, 6$).

Sequence of information acquisition. The sequence of information acquisition (SIA) on each attribute was defined operationally as the average rank order of acquiring information on that dimension across the eight records:

$$SIA_j = \sum_{i=1}^8 RANKACQ_{ij} / 8 \quad (2)$$

where $RANKACQ_{ij}$ is either the rank order in which a cue was actually selected in the two tasks combined or the average rank order for those remaining cues not selected in either task.

TABLE 1

Matrix of Information Representing Record Albums
Specified According to a Fractional Factorial Design

	1. Singer's Style	2. Type of Songs	3. Type of Production	4. Jacket	5. Label	6. Price
Record A	contemporary	standard	live	visual	major	\$5.59
Record B	contemporary	recent	studio	informative	independent	\$5.59
Record C	contemporary	standard	live	informative	independent	\$4.19
Record D	traditional	standard	studio	informative	major	\$5.59
Record E	contemporary	recent	studio	visual	major	\$4.19
Record F	traditional	recent	live	visual	independent	\$5.59
Record G	traditional	recent	live	informative	major	\$4.19
Record H	traditional	standard	studio	visual	independent	\$4.19

Information value. Several alternative indices of information value were examined for their ability to predict the extent and sequence of information acquisition. In addition to the attribute-specific scores for importance (IMP_j), confidence (CON_j), and uncertainty (UNC_j), multiplicative information-value indices were defined as follows:

$$\text{importance-confidence index} = IMP_j \times CON_j \quad (3)$$

$$\text{importance-uncertainty index} = IMP_j \times UNC_j \quad (4)$$

Attitude structure. In accord with the multiplicity of available multiattribute models of attitude structure (Bass and Wilkie, 1973; Fishbein and Ajzen, 1975; Holbrook, 1976; Wilkie and Pessemier, 1973), several alternative attitude models were examined for their ability to predict preference order:

$$\text{Partial Additive Model} = \sum_{k=1}^6 B_k \cdot E_k \cdot ACQ_k \quad (5)$$

$$\text{Partial Averaging Model} = \left(\sum_{k=1}^6 B_k \cdot E_k \cdot ACQ_k \right) / \left(\sum_{k=1}^6 ACQ_k \right) \quad (6)$$

$$\text{Full Additive Model} = \sum_{k=1}^{12} B_k \cdot E_k \cdot ACQ_k \quad (7)$$

$$\text{Full Averaging Model} = \left(\sum_{k=1}^{12} B_k \cdot E_k \cdot ACQ_k \right) / \left(\sum_{k=1}^{12} ACQ_k \right) \quad (8)$$

where k refers to the twelve characteristics that a given record does ($k = 1, \dots, 6$) and does not ($k = 7, \dots, 12$) possess; $B_k = 1$ if the record possesses characteristic k or $B_k = -1$ if it does not; E_k represents a respondent's evaluation of the desirability of characteristic k (scored from -3 to $+3$ or from 1 to 7); and ACQ_k is a zero-one dummy variable representing the acquisition of information on characteristic k .

In addition to the attitude models listed above, importance-weighted versions were also investigated. These were identical to the above forms except that $B_k \cdot E_k \cdot ACQ_k$ was replaced by $B_k \cdot E_k \cdot IMP_k \cdot ACQ_k$. Note that, for all such models, the inclusion of the multiplicative ACQ_k term has the effect of setting belief scores equal to zero for those characteristics about which information was not acquired. Note also that the Partial Models assume that preference depends only upon the evaluations

of the characteristics known to be present in a record whereas the Full Models assume that preference may be enhanced (reduced) by the absence of undesirable (desirable) characteristics. The Full Models thus incorporate logic analogous to that underlying arguments for the bipolar coding of belief and evaluative-aspect scores in the Fishbein model (Fishbein, 1976; Fishbein and Ajzen, 1975).

Hypotheses

Given the operational measures defined above, two key hypotheses were tested:

H_1 . Information value predicts the extent and sequence of information acquisition

H_2 . Attitude-structure models based on the information acquired predict preference rank

There were no a priori hypotheses concerning the relative predictive efficacy of importance, confidence, uncertainty, the importance-confidence index, and the importance-uncertainty index as measures of information value or of the Partial and Full Additive and Averaging Models as representations of attitude structure. Rather, findings concerning the relative performance of these alternative indices and models were considered exploratory.

Statistical Procedures

Statistical routines were run separately for each respondent using (1) Pearson product-moment correlations across the six attributes to test hypothesis H_1 and (2) Spearman rank-order correlations across the eight records to test hypothesis H_2 . The use of intraindividual analysis is considered essential in such correlational tests to avoid utility comparisons between individuals and to minimize yea-saying bias (Bass and Wilkie, 1973).

Results

Few problems appeared to arise in the respondents' self-administration of the instrument described above. All but a handful of the questionnaires were returned in proper order (the exceptions were, of course, discarded), and where a sticker was occasionally damaged or defaced, respondents showed a reassuring tendency to write in the

TABLE 2

Mean and Median Pearson Correlations Within Individuals
Between Measures of Information Value
and Indices of Extent and Sequence of Information Acquisition

Measure of Information Value	Index of Information Acquisition					
	Extent			Sequence		
	Mean Correlation ^a	Standard Error ^a	Median Correlation ^a	Mean Correlation ^a	Standard Error ^a	Median Correlation ^a
Importance	.587	.038	.722	-.633	.037	-.752
Confidence	.186	.043	.056	-.161	.045	-.097
Uncertainty	.055	.062	.000	-.052	.064	-.000
Importance-Confidence Index	.543	.040	.701	-.577	.040	-.727
Importance-Uncertainty Index	.438	.046	.609	-.478	.047	-.703

^aN = 100.

missing identification number by hand or to use a piece of Scotch tape to stick it firmly in place.

Results for hypothesis H₁ are presented in Table 2. Attribute importance offered fairly good mean predictions of both extent and sequence of information acquisition ($\bar{r} = .587$, $Z_{\bar{r}} = 15.45$, $p < .0001$; $\bar{r} = -.633$, $Z_{\bar{r}} = -17.11$, $p < .0001$). Because the distributions of these correlations were skewed, the medians were substantially stronger than the means ($\bar{r} = .722$ vs. $.587$; $\bar{r} = -.752$ vs. $-.633$). The predictive performance of attribute importance was significantly better than that of either confidence or uncertainty: $\bar{r} = .587$ vs. $.186$ ($t_{99} = 8.33$, $p < .0001$); $\bar{r} = .587$ vs. $.055$ ($t_{99} = 7.38$, $p < .0001$); $\bar{r} = -.633$ vs. $-.161$ ($t_{99} = 9.25$, $p < .0001$); $\bar{r} = -.633$ vs. $-.052$ ($t_{99} = 8.18$, $p < .0001$). Finally, the multiplicative indices of information value (importance-confidence and importance-uncertainty) failed to improve upon the predictions obtained by importance considered alone: $\bar{r} = .587$ vs. $.543$ ($t_{99} = 2.87$, $p < .01$); $\bar{r} = .587$ vs. $.438$ ($t_{99} = 3.82$, $p < .001$); $\bar{r} = -.633$ vs. $-.577$ ($t_{99} = 3.11$, $p < .01$); $\bar{r} = -.633$ vs. $-.478$ ($t_{99} = 3.90$, $p < .001$). In sum, among those variables tested, attribute importance (with median correlations above $\bar{r} = .70$) was the simplest and best index of information value for predicting the extent and sequence of information seeking.

Table 3 presents the results for hypothesis H₂. In no case did the coding of desirability (E_k) from 1 to 7, instead of from -3 to +3, have any statistically significant impact on the performance of the Partial or Full Additive or Averaging Models in predicting preference. Nor did the inclusion of importance weights provide any significant improvement in the predictive performance of these models (probably because the models included only those belief scores considered important enough to have prompted information seeking in the first place). Table 3 therefore contains only the rank-order correlational results for the simple Partial and Full Additive and Averaging Models with desirability (E_k) coded in the theoretically appropriate manner,

TABLE 3

Mean and Median Spearman Rank-Order Correlations
Within Individuals Between Brand Preference
and Models of Attitude Structure

Model of Attitude Structure	Mean Rank-Order Correlation ^a	Standard Error ^a	Median Rank-Order Correlation ^a
Partial Additive Model ^b	-.568	.039	-.704
Partial Averaging Model ^b	-.542	.041	-.670
Full Additive Model ^b	-.558	.040	-.699
Full Averaging Model ^b	-.551	.040	-.683

^aN = 94 due to some cases with missing data.

^bIn all cases, results are reported for the models without importance weights and with desirability (E_k) coded from -3 to +3. These results did not differ significantly in any way from those obtained with importance-weighted models or with a 1-to-7 coding of E_k .

from -3 to +3. It is clear that all four versions of the model supported hypothesis H₂ by performing fairly well, with mean rank-order correlations ranging from $-.542$ to $-.568$ ($Z_{\bar{r}} = 13.22$, $p < .0001$). Again, the distributions of these correlations are skewed so that the medians are considerably stronger, ranging from $-.670$ to $-.704$. Since there is no statistically significant difference between the correlations obtained by any pair of

these four models, it might be argued that the "best" is therefore the simplest--namely, the Partial Additive Model, which simply sums the desirability scores of those characteristics known (after information seeking) to be possessed by a particular record. Although these results concerning the relative performance of the Additive and Averaging Models were considered exploratory in the present research, this finding may reinforce doubts about the often-claimed superiority of the averaging formulation (Lutz, 1976; cf. Fishbein, 1976; Fishbein and Ajzen, 1975). Further work comparing additive and averaging versions, using the sticker-removal task, is in progress.

Discussion

Limitations

The study reported above is subject to several important limitations. First, though an effort was made to select a product (phonograph records) of considerable relevance to the student population, it may be that, in acquiring information to evaluate records, students behave differently from housewives evaluating other consumer products (e.g., toothpaste), thus limiting the generalizability of the findings. Moreover, the records used were identified only by letter rather than by title and name of artist. This procedure helped guard against confounding effects, but represented a further departure from the real market place. It might also be objected that the self-administration of questionnaires passed out by hand differs in some nonobvious way from the completion of those received through the mail. Given the relative anonymity of the respondents, however, this potential problem does not appear to have been too damaging in the present situation.

Perhaps most seriously, the study's design may be subject to the impact of task structure on acquisition strategy recently demonstrated by Bettman and Kakkar (1977). For example, the first attribute (singer's style) was widely regarded as the most important, thus suggesting the alternative hypothesis that the relationship between importance and acquisition may have resulted, in part, from a tendency for respondents to begin at the left of the matrix and work their way to the right. This interpretation is at least partly ruled out by the fact that the last attribute (price), on the far right, was also generally regarded as highly important. Nevertheless, future work with the sticker-removal task will attempt to eliminate such potential order effects.

Conclusions

Apart from these limitations, the study seems to have achieved its purpose of using a new technique, suitable for administration by survey questionnaire, to test two key hypotheses concerning the relation of information acquisition to the multiattribute model. In general, the findings support the key role played by attribute importance at the interface between information acquisition and attitude structure. Specifically, in accord with hypothesis H_1 , attribute importance appears to guide search toward the most important attribute-related cues. Further, in accord with hypothesis H_2 , these most important cues appear to determine preference rank. This support for hypotheses H_1 and H_2 suggests a determinant role of perceived attribute importance in directing search toward information that, in turn, shapes the beliefs incorporated into attitude structure.

Finally, the study demonstrated the usefulness of the questionnaire-based sticker-removal task for investigating information acquisition. If consumer researchers are to heed Ferber's (1977) recent call for studies

using more representative samples, with a probable resultant loss in the feasibility of applying the techniques of laboratory experimentation, some such survey approach to studying information acquisition may well be necessary in the future.

References

- O. T. Ahtola, "The Vector Model of Preferences: An Alternative to the Fishbein Model," Journal of Marketing Research, 12(February, 1975), 52-59.
- F. M. Bass and W. L. Wilkie, "A Comparative Analysis of Attitudinal Predictions of Brand Preference," Journal of Marketing Research, 10(August, 1973), 262-269.
- D. E. Berlyne, Conflict, Arousal, and Curiosity (New York: McGraw-Hill, 1960).
- J. R. Bettman, "Toward a Statistics for Consumer Decision Net Models," Journal of Consumer Research, 1(June, 1974), 71-80.
- J. R. Bettman, "Data Collection and Analysis Approaches for Studying Consumer Information Processing," Working Paper No. 41, Center for Marketing Studies, University of California, Los Angeles, 1976.
- J. R. Bettman and J. Jacoby, "Patterns of Processing in Consumer Information Acquisition," Proceedings, Association for Consumer Research, Sixth Annual Conference, 1976, 315-319.
- J. R. Bettman and P. Kakkar, "Effects of Information Presentation Format on Consumer Information Acquisition Strategies," Journal of Consumer Research, 3(March, 1977), 233-240.
- R. W. Chestnut, "The Impact of Energy-Efficiency Ratings: Selective Vs. Elaborative Encoding," Purdue Paper in Consumer Psychology No. 160, Purdue University, 1976.
- D. F. Cox, ed., Risk Taking and Information Handling in Consumer Behavior (Cambridge, Mass.: Harvard University Graduate School of Business Administration, 1967).
- W. Edwards, "Optimal Strategies for Seeking Information: Models for Statistics, Choice Reaction Times, and Human Information Processing," Journal of Mathematical Psychology, 2(1965), 312-329.
- R. Ferber, "Research By Convenience," Journal of Consumer Research, 4(June, 1977), 57-58.
- M. Fishbein, "Extending the Extended Model," Proceedings, Association for Consumer Research, Sixth Annual Conference, 1976, 491-497.
- M. Fishbein and I. Ajzen, Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research (Reading, Mass.: Addison-Wesley, 1975).
- M. Glassman and N. Fitzhenry, "Fishbein's Subjective Norm: Theoretical Considerations and Empirical Evidence," Proceedings, Association for Consumer Research, Sixth Annual Conference, 1976, 477-480.
- P. E. Green, "Uncertainty, Information, and Marketing Decisions," in R. Cox, W. Alderson, and S. J. Shapiro, eds., Theory in Marketing (Homewood, Ill.: Irwin, 1964), 333-354.
- P. E. Green, "On the Design of Choice Experiments Involving Multifactor Alternatives," Journal of Consumer Research, 1(September, 1974), 61-68.

- G. H. Haines, Jr., "Process Models of Consumer Decision Making," in G. D. Hughes and M. L. Ray, eds., Buyer/Consumer Information Processing (Chapel Hill: University of North Carolina Press, 1974), 89-107.
- M. B. Holbrook, "A Study of Communication in Advertising," unpublished doctoral dissertation, Columbia University, 1975.
- M. B. Holbrook, "Attitude Structure and Communication Effects: The Neglected Interface," working paper, Columbia University, 1976.
- M. B. Holbrook, "Beyond Attitude Structure: Toward the Informational Determinants of Attitude," working paper, Columbia University, 1977.
- M. B. Holbrook and J. M. Hulbert, "Multi-Attribute Attitude Models: A Comparative Analysis," Proceedings, Association for Consumer Research, Fifth Annual Conference, 1975, 375-388.
- J. Jacoby, "Perspectives on a Consumer Information Processing Research Program," Communication Research, 2(July, 1975), 203-215.
- J. Jacoby, "The Emerging Behavioral Process Technology in Consumer Decision-Making Research," Purdue Paper in Consumer Psychology No. 169, Purdue University, 1976.
- J. Jacoby, R. W. Chestnut, and W. A. Fisher, "Simulating Nondurable Purchases: Individual Differences and Information Acquisition Behavior," unpublished paper, Purdue University, n.d.
- J. Jacoby, R. W. Chestnut, K. C. Weigl, and W. A. Fisher, "Pre-Purchase Information Acquisition: Description of a Process Methodology, Research Paradigm, and Pilot Investigation," Proceedings, Association for Consumer Research, Sixth Annual Conference, 1976, 306-314.
- J. Jacoby, G. J. Szybillo, and J. Busato-Schach, "Information Acquisition Behavior in Brand Choice Situations," Journal of Consumer Research, 3(March, 1977), 209-216.
- R. J. Lutz, "Conceptual and Operational Issues in the Extended Fishbein Model," Proceedings, Association for Consumer Research, Sixth Annual Conference, 1976, 469-476.
- J. G. Myers, "An Interactive Computer Approach to Product Positioning," paper presented at the Attitude Research Conference, Hilton Head, South Carolina, February, 1976.
- M. Nakanishi and J. R. Bettman, "Attitude Models Revisited: An Individual Level Analysis," Journal of Consumer Research, 1(December, 1974), 16-21.
- J. W. Payne, "Heuristic Search Processes in Decision Making," Proceedings, Association for Consumer Research, Sixth Annual Conference, 1976a, 321-327.
- J. W. Payne, "Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis," Organizational Behavior and Human Performance, 16(1976b), 366-387.
- I. Ross, "Perceived Risk and Consumer Behavior: A Critical Review," Proceedings, Association for Consumer Research, Fifth Annual Conference, 1975, 1-19.
- J. W. Taylor, "The Role of Risk in Consumer Behavior," Journal of Marketing, 38(April, 1974), 54-60.
- D. J. Tigert, "An Investigation of Experimental Techniques to Determine the Relative Importance of Various Product Attributes in the Consumer Decision-Making Process," Institute Paper No. 126, Institute for Research in the Behavioral, Economic, and Management Sciences, Krannert Graduate School of Industrial Administration, Purdue University, 1966.
- W. L. Wilkie and E. A. Pessemier, "Issues in Marketing's Use of Multi-Attribute Attitude Models," Journal of Marketing Research, 10(November, 1973), 428-441.
- F. W. Winter, "Laboratory Measurement of Response to Consumer Information," Journal of Marketing Research, 12(November, 1975), 390-401.
- P. L. Wright, "The Cognitive Processes Mediating Acceptance of Advertising," Journal of Marketing Research, 10(February, 1973), 53-62.
- P. L. Wright, "The Harassed Decision Maker: Time Pressures, Distractions, and the Use of Evidence," Journal of Applied Psychology, 59(1974), 555-561.
- P. L. Wright, "Consumer Choice Strategies: Simplifying Vs. Optimizing," Journal of Marketing Research, 12(February, 1975), 60-67.

ALTERNATIVE APPROACHES TO ASSESSING
THE QUALITY OF SELF REPORT DATA

Robert A. Hansen, The Ohio State University
Carol A. Scott, The Ohio State University

Abstract

Consumer behavior researchers place a great deal of emphasis on self-report or questionnaire data as input to decision making, but give little attention to assessing the quality of subjects' responses. A review of classical reliability theory indicates that available reliability measures are inappropriate in many consumer behavior data collection situations. This is particularly true in those cases where non-scaled or ad hoc measures are collected. Data quality requirements are discussed and some alternative methods of assessing data quality are presented.

Introduction

Consumer behavior researchers depend heavily on self-report data generated through mail survey questionnaires or interview schedules for information about consumer attitudes, preferences, and behavior. Self-report data is an integral part of laboratory and field experiments as well as survey research, and the recent explosion of articles on generating responses to self-report instruments attests to its importance (see Houston and Ford, 1976; Kanuk and Berenson, 1975). Pressley, 1976 for reviews). Much of this work, however, has been limited to a focus on the quantity of data collected primarily through mail questionnaires. Little attention has been devoted to the parallel issue of the quality of self-report data regardless of the collection method used.

This paper attempts to examine the topic of self-report data quality in some detail. Specifically, the objectives of this paper are to critically evaluate current methods of assessing data quality in those situations where traditional methods cannot be used. To provide an adequate context for this discussion, basic data quality requirements are first reviewed.

Data Quality Requirements

At least two broad criteria for data quality can be identified. First, responses must be meaningful. That is, they should reflect something other than a random checking of alternatives. This implies that responses reflect a true assessment of a particular behavior (e.g., frequency of purchase) or a true assessment of some personal characteristic (e.g., attitudes, brand loyalty, dogmatism), as opposed to unsystematic, situation-specific variation. The issue here is typically thought of as the degree of reliability of the responses.

Second, responses should be good indicators of the phenomenon of interest. That is, given that an adequate degree of reliability has been achieved, one must demonstrate that the scores are an adequate representation of the construct or quality under investigation. The issue in this case is one of validity (predictive, content, construct, etc.), which is examined by investigating the sources of the systematic variation observed.

This paper will focus primarily on the first of these criteria since reliability is necessary (but not sufficient) for validity. A second reason for this emphasis is that there are unique problems associated with the assessing the meaningfulness of responses to nonscaled measures which are frequently used by consumer behavior researchers and other applied scientists, and these

problems have received relatively little attention in the consumer behavior literature. Finally, the soundness of decisions (whether of managerial or public policy significance) based upon research data is profoundly affected by the degree of reliability associated with the data. As one prominent methodologist notes (Stanley, 1971, p. 58):

Unreliability places a question mark after the score (or response) and causes any judgment based on it to be tentative to some extent. The lower the reliability of the score, the more tentative the judgment or decision must be, until, in the extreme case, as reliability approaches zero, the score (or response) provides no basis at all for any judgment or decision...

Traditional Methods of Assessing Reliability

Appropriate methods of assessing the reliability of data depend upon the type of measurement instrument used. Consumer behavior researchers generally use two types of instruments: scaled measures and non-scaled, or ad hoc, measures. A scaled measure is defined here to be the result of pooling or combining a number of items selected from a larger pool to yield a score on the particular dimension of interest (e.g., androgeny, authoritarianism, attitudes toward government, etc.). Non-scaled or adhoc measures are those which are not designed to be combined for a summary score, but rather are of interest in and of themselves (e.g., frequency of purchase, preference for a given brand, etc.). Although specific methods for assessing reliability will differ depending upon the type of measurement instrument used, the logic should be similar. That is, one should gain some insight into the degree of systematic variation captured by the instrument. The discussion here is an evaluation of the appropriateness of currently available techniques for scaled and non-scaled data.

Scaled Measures

When the measurement instrument consists of a series of items which form a scale, assessing reliability can be accomplished by using one of several traditional methods. The mechanics of these techniques are outlined, and the relative merits of each are discussed, in several marketing research textbooks (e.g., Churchill, 1976; Tull and Albaum, 1973; Tull and Hawkins, 1976) and in several recently published articles (e.g., Lundstrom and Lamont, 1976; Pressley, 1976).

If two measurements are possible, a test-retest correlation can be calculated. This is perhaps not the best alternative, however. It is impractical in many consumer behavior research contexts to contact respondents twice for the necessary repetition of the measure. And, many constructs of interest to consumer researchers are variable over time. Since attitudes toward brands can change substantially even in the short run, for example, a low test-retest correlation would be difficult to interpret. It might be possible to administer parallel forms of the same instrument, but this may be possible only in those situations in which a researcher or his representative is present to provide an explanation for any apparent repetition, and where length of time

necessary to complete the parallel forms is not prohibitive.

Equivalence or internal consistency measures are also appropriate for a set of items which form a scale for which a summary score will be derived (Cronbach, 1951; Kuder and Richardson, 1937). This type of coefficient, unlike test-retest, is easily obtained, and should be calculated routinely for scaled instruments (Nunnally, 1967). A high reliability coefficient will provide some evidence that the responses reflect a systematic tendency, and will provide some rationale for combining the responses to form a scale score. Unlike test-retest procedures, equivalence coefficients should not pose any practical difficulties in terms of supervisory or time needs.

In summary, researchers should follow the standard procedures outlined in measurement texts in those instances where scaled measures are used. In some situations, a scaled measure does not exist, but could be constructed. That is, the concept is potentially scalable, but consumer researchers have not yet attempted to do the scale development work. It should be clear, however, that scales should be constructed whenever possible (Hughes, 1971; Lundstrom and Lamont, 1976).

Ad Hoc Measures

In some situations, the researcher cannot or perhaps would not want to construct a scaled measure, and a non-scaled or ad hoc measure is used instead. These non-scaled measures present some difficulty for quality assessment. The following section discusses ad hoc independent and dependent measures in greater detail.

Ad Hoc Independent Measures. In assessing the reliability of ad hoc independent or predictor measures, the use for which the data is intended must be considered since some procedures will be feasible for theory-testing research and others will be appropriate for applied or problem-solving research. An example using the multi-attribute product evaluation framework will illustrate this difference. If a researcher is interested in predicting product preference or purchase behavior, a strong case can be made for using the domain sampling model (Nunnally, 1967) and selecting multiple items which represent all aspects of the product evaluation process. Traditional methods such as Cronbach's coefficient alpha could then be used to assess reliability via internal consistency. An applied researcher, however, is frequently interested in evaluating a select group of the product's attributes. Further, he is most likely interested in the individual attributes themselves rather than in an overall summary score. Test-retest procedures are generally not acceptable for practical reasons, and internal consistency is not appropriate since a summary score is not desired. As yet, there are no widely accepted methods for assessing the reliability of this type of data.

Ad Hoc Dependent Measures. Criterion measures such as purchase behavior or product usage must also be examined from a reliability standpoint since reliable predictor variables applied to an unreliable criterion may produce misleading results. Many criterion variables are analogous to physical behavior and motor skills measures discussed by (Stanley, 1971). That is, an item pool does not exist because there are only a small number of ways in which certain questions can be asked. Since these variables are not scalable, traditional measures of reliability cannot be used as indicators of data quality, and therefore, other methods must be developed.

Alternative Approaches to Examining Data Quality

There are several techniques which the consumer behavior researcher can use to infer data quality when traditional

tools are not appropriate. These methods include analysis of the completeness of response, depth of response, and internal consistency of response patterns. These methods are admittedly non-traditional with only internal consistency of response patterns bearing a slight resemblance in title and operation to a psychometric interpretation of reliability. It must be remembered, however, that the data collection needs of the consumer behavior researcher are also quite often non-traditional in nature requiring the collection of ad hoc or non-scaled data. The results of these analyses can be used in two ways. First, the researcher who has collected data from a number of people can identify those individuals whose responses might be of questionable quality. These questionable respondents might then be discarded from the data base. Second, they may be used to assess the overall effectiveness of various methods used to collect data. Viewed in this manner, the measures could be used, for example, to evaluate the effects of alternative methods of stimulating mail survey response rates (Houston and Ford, 1976).

Completeness of Response

Completeness deals with the issue of item non-response. Of course the researcher's goal is to generate 100 percent response to all questions. Any deviation from total response to all questions raises a number of difficult problems. In many instances, methods used to handle missing values (e.g., drop those cases from the analysis, substitute a mean value, etc.) can have a serious effect in the representativeness of the results (Hansen and Scott, 1977).

Item-nonresponse can be expressed in a variety of ways. These include the mean number of missing values, the percentage of missing values, and complete versus incomplete response (Houston and Ford, 1976). However non-response is defined, any modification of the data collection process which yields a lower item non-response is desirable. Modification of a mail survey could mean the inclusion of an incentive to respond or the promise of anonymity. A second modification of the data collection process could involve the researcher allowing respondents to ask questions or by providing a don't know option for questions.

Depth of Response

Quality can also be viewed in terms of response to open-ended questions. The additional insight provided by an analyses of the depth of response is illustrated by the following example from a project currently being completed by one of the authors. In a mail survey, potential respondents were randomly assigned to one of three incentive groups: a monetary incentive group (i.e., 25¢ included with the survey), a non-monetary incentive group (i.e., a comparable value ball point pen included with the survey), and a no incentive group. The response rates for the three treatments (which indicate the quantity of data) were 38%, 22% and 14%, respectively. These results seem to indicate rather dramatically the effect of including inducements in the survey. However, when the returned questionnaires were analyzed for percent of missing values, the rates were significantly higher for both incentive groups. In addition, an independent panel of three judges rated the responses to a series of open-ended questions on the questionnaire and these ratings were subsequently analyzed. Results indicated that a significantly greater proportion of the responses in the non-incentive group were rated superior by the judges. In general the overall quality of response from the non-incentive control group was significantly better than that for the two incentive groups. Clearly in this case at least, the quality of data obtained issue causes us to re-evaluate the supposed superior performance of the inducements to respond to a mail survey.

There is another important aspect of assessing data quality which relates to those questionnaires where the respondent has in fact answered all or nearly all questions and where content analysis of responses is not possible. One would, of course, suspect the quality of responses from a person whose questionnaire contained a significant number of nonanswered questions. The equally important but more complex issue of evaluating the quality of a completed questionnaire is discussed below.

Internal Consistency

Internal consistency as used here is perhaps most closely associated with the concept of response bias. Response bias is frequently mentioned in connection with the self selection problem in mail surveys. In a number of these studies mathematical estimates of bias have been generated. (Kanuk and Berenson, 1975) discuss this interpretation of bias and the controversy surrounding its use. In addition response bias has been investigated by validating responses given in self-report sessions against other records of the same information (Kerin, 1974; Kerin and Peterson, 1977). In still other cases, bias has been defined as a difference in response pattern derived from a similar population using for example, different incentives to respond (Whitmore, 1976) or other attempts to stimulate response (Field, 1975; Wiseman, 1972). Bias it appears then has been used as a general umbrella to encompass a number of processes from true validity checks to simple checks of difference. In the former, it is assumed that one response is true (and deviations are errors), and in the latter it is just assumed that differences are bad. Bias, when operationalized using a known or accepted as true validating test, captures the essence of the data quality issue. Bias when defined in this manner focuses on the question of whether the responses are indicative of a true feeling as opposed to a random response or systematically altered response.

There are, however, two problems associated with using validated response bias as an overall indicator of data quality. First, there is the practical problem of collecting the validation information. This kind of information is only rarely available (Kerin, 1974; Kerin and Peterson, 1977) and if it has to be collected in addition to the data of interest it will increase project costs considerably.

A second drawback associated with using bias as a quality indicator is offered by Ferber (1948, p. 670): "The problem of response bias must be considered with specific reference to a particular question or characteristic. The presence of bias in one question does not mean a priori that the replies to other questions on the same questionnaire are also biased."

This suggests that question dependent response bias is not an acceptable indicator of overall response quality. What is needed is an assessment of quality derived from the entire self report data collection process. One possibility for assessing data quality which seems consistent with this goal could include bogus questions (Friedman and Goldstein, 1975), but could be expanded to include simple response inconsistencies. Additional questions may have to be added to the data collection instrument, but in many cases this will not be necessary. Careful attention to question wording can be used to generate logic checks. In this way the overall quality of respondent answers is inferred by the overall logic or consistency in response patterns. For example, if a respondent indicates that he/she has never heard of a particular product and then proceeds to evaluate that product, then one should be suspicious. Or, sequential sets of questions may be developed to allow for examination of the pattern of responses. If a respondent indicates that he/she never engages in a behavior and then in a later section answers a frequency question with a

response of greater than zero, one must be suspicious of the quality of the data.

Clearly, there are a number of safeguards which must be built into the analysis of such logic checks. First and foremost, all procedures must be pre-tested for clarity of instructions and ease of understanding. Second, one must examine the incidence of logic errors. One such error should not be taken as evidence of poor quality or response. Several errors of this type, however, would indicate problems with the quality of responses.

An example should help illustrate the usefulness of the process in detecting problems with data collected. A study was conducted by one of the authors in which a number of logic checks were built into a series of questions dealing with attitudes toward receiving direct mail, readership habits with regard to direct mail, and actions taken in response to direct mail (e.g., use of coupons, shop at store, sales, etc.). Results of the study indicated that approximately six percent of the returned questionnaires contained some logic errors. Only a small proportion of these poor quality respondents had a high proportion of missing values. If it is assumed that the incidence of logic checks indicates poor quality, then attempting to "clean-up" the data by setting some cut off point for an acceptable level of missing values would not accomplish the goal. Low data quality can be a problem even when item non-response is low.

Discussion and Conclusion

The quality of self-report data is of central importance to consumer behavior researchers. As shown in this paper, however, the applied researcher must develop his own indices of response quality in many instances. Where scaled measures exist or can be developed, standard measures of data quality exist and should be used. In those cases where these techniques are not applicable, data quality cannot be ignored. Several checks on the quality of responses have been suggested in this paper, and researchers can doubtlessly think of many others.

The procedures suggested are stop gap measures at best. Clearly, more work is needed in this area to develop standard techniques of assessing the quality of responses to non-scaled measures. Research is needed which has as its primary focus the quality of self-report data. This research should attempt to investigate definitions and measures of data quality and their relationship to standard definitions and measures of reliability and validity. Once these measures and definitions are developed, replication of their operationalization must be carried out across topics and populations.

A second area of research should focus on methods of increasing the quality of data. Instructions, for example, may be tested for their ability to provide better data. The inclusion of a "don't know" option should be tested to determine whether it will reduce item non-response and inconsistent patterns of data. Or, special questions may be included to screen out those respondents whose opinions and answers may be unreliable. For example, respondents could be asked to indicate how certain they are of their judgments. Respondents who are extremely uncertain of their answers may be dropped from further analyses. The efficacy of this approach could be tested by examining the correlation between certainty scores and test-retest reliability scores. Finally, one may conclude that special questionnaires must be developed for different populations of respondents who demonstrate difficulty with standard questionnaire formats.

In any case, it is not enough to suggest that researchers develop scales for every measurement situation and use

traditional reliability and validity techniques. There are too many instances in applied research settings where variables are not scalable, or where a score is not particularly useful to the researchers. A better tactic is to recognize the issue of data quality and to develop measures to fit the research context.

References

- Gilbert A. Churchill, Jr., Marketing Research: Methodological Foundations (Homewood, Illinois: Dryden Press, 1976).
- Lee J. Cronbach, "Coefficient Alpha and the Internal Structure of Tests," Psychometrika 16(June, 1951), 297-334.
- Hubert S. Field, "Effect of Sex of Investigator on Mail Survey Response Rates and Response Bias," Journal of Applied Psychology, 60(December, 1975), 772-773.
- Robert Ferber, "The Problem of Bias in Mail Returns: A Solution," Public Opinion Quarterly, 12(Winter, 1948), 669-676.
- Hershey H. Friedman and Larry Goldstein, "Effect of Ethnicity of Signature on Rate of Return and Content of a Mail Questionnaire," Journal of Applied Psychology, 60(December, 1975), 770-771.
- Robert A. Hansen and Carol A. Scott, "Improving the Representativeness of Survey Research: Some Issues and Unanswered Questions," Educator's Proceedings, American Marketing Association, 1977, 401-4.
- Michael J. Houston and Neil M. Ford, "Broadening the Scope of Methodological Research on Mail Surveys," Journal of Marketing Research, 13(November, 1976), 397-403.
- David G. Hughes, Attitude Measurement for Marketing Strategies, (Glenview, Illinois: Scott Foresman and Company, 1971).
- Leslie Kanuk and Conrad Berenson, "Mail Surveys and Response Rates: A Literature Review," Journal of Marketing Research, 12(November, 1975), 440-452.
- Roger A. Kerin, "Personalization Strategies, Response Rate and Response Quality in a Mail Survey," Social Science Quarterly, 55(June, 1974), 175-181.
- Roger A. Kerin and Robert A. Peterson, "Personalization, Respondent Anonymity and Response Distortion in Mail Surveys," Journal of Applied Psychology, 62(November, 1977), 86-89.
- G. F. Kuder and M. W. Richardson, "The Theory of the Estimation of Test Reliability," Psychometrika, 2 (September, 1937), 151-160.
- Arnold S. Linsky, "Stimulating Responses to Mailed Questionnaires: A Review," Public Opinion Quarterly, 39(Spring, 1975), 82-101.
- William J. Lundstrom and Lawrence M. Lamont, "The Development of a Scale to Measure Consumer Discontent," Journal of Marketing Research, 13(November, 1976), 373-381.
- J. C. Nunnally, Psychometric Theory, (New York: McGraw-Hill, 1967).
- Paul J. Peter, "Reliability Generalizability and Consumer Behavior," Proceedings, Association for Consumer Research, 1976, 394-400.
- Milton M. Pressley, Mail Survey Response: A Critically Annotated Bibliography, (Greensboro, North Carolina: Faber and Company, 1976).
- Julian Stanley, "Reliability," in Robert L. Thorndike (ed.), Educational Measurement, 2nd ed., (Washington, D.C.: American Council on Education, 1971), 356-442.
- Donald S. Tull and Gerald S. Albaum, Survey Research: A Decisional Approach, (New York: Intex Educational Publishers, 1973).
- Donald S. Tull and Del I. Hawkins, Marketing Research: Meaning, Measurement, and Method, (New York: MacMillan Publishing Company, 1976).
- William J. Whitmore, "Mail Survey Premiums and Response Bias," Journal of Marketing Research, 13(February, 1976), 46-50.
- Frederick Wiseman, "Methodological Bias in Public Opinion Surveys," Public Opinion Quarterly, 36(Spring, 1972), 105-108.

AN EVALUATION OF TELEPHONE SAMPLING DESIGNS

E. Laird Landon, Jr., University of Houston
Sharon K. Banks, University of Oregon

Introduction

The use of telephone interviewing in survey research has increased rapidly in the past decade. The increasing costs of personal interviewing, coupled with the higher crime rate in inner cities, have made personal interviewing less practical. On the other hand, the inherent biases and problems assumed to be created by the use of telephones in surveys appears to have been reduced. The percentage of households with telephones has reached an extremely high level. Improved telephone services, including WATS and pushbutton phones have increased efficiency. Sampling methods have been improved to better reflect the populations being sampled. These factors have contributed to the tendency of many researchers to substitute telephone interviewing for personal interviewing where appropriate. When considering the use of telephone interviewing, the adequacy of the sampling frame is of vital concern to the researcher.

Directory Sample Frame

Any research method which exclusively uses the telephone regardless of sample design may produce biased estimates because of the exclusion of non-telephone households. The extent of the bias depends upon how different households without telephones are, with respect to the issues under study. Studies conducted in 1968 and 1969 in Missouri showed that households without telephones tended to have lower incomes and live in rural areas more often than did households with telephones (Leuthold and Scheele, 1971). On the other hand, Klecka and Tuchfarber (1975b) state that, "... even personal interviewing techniques have difficulty locating the very people who are least likely to have phone service--namely, transients, the exceptionally disadvantaged and social dropouts." Also, the researcher may be less interested in the responses of non-telephone households. Because of the research objectives any bias may be unimportant.

Additionally, the percentage of households without telephones is apparently growing smaller. By 1976 over 90 percent of the households in the United States could be reached by telephone (Tuchfarber, et al., 1976). Sudman (1973) estimated that 98.5 percent of households in two Chicago suburbs had telephones.

Sample designs based only on telephone directories may suffer additional biases due to the many households with telephones which are not listed. Some households do not want to be in the directory so they are voluntarily unlisted--often paying a charge for the service. However, many households are not in the directory for reasons beyond their control; these households are involuntarily unlisted. Households may be involuntarily unlisted because they have recently moved into the community, they have moved within the community, for some other reason their number has been changed, or an error has occurred in producing the directory (Cooper, 1964). The total percentage of non-listed households has been estimated to be between 18-20 percent nationally (Cooper, 1964; Glasser and Metzger, 1975). However, unlisted percentages were estimated to be as high as 29 percent in some large metropolitan counties (Glasser and Metzger, 1975) and below 5 percent in some rural areas (Sudman, 1973).

The age of the directory is also a contributing factor to the percentage of unlisted households. A new direc-

tory may exclude only about 3 percent of all households involuntarily, while a year-old directory might exclude as many as 12 percent (Perry, 1968-69). The voluntarily unlisted percentage is probably fairly stable throughout the year--perhaps between 6 percent and 13 percent (Brunner and Brunner, 1971).

While the exact nature of bias created by not listing certain households depends on the population and the topic under investigation, some studies have discovered significant differences in unlisted households. In a 1967 study in Toledo, Ohio, Brunner and Brunner (1971) found that heads of voluntarily unlisted households tended to have less education, tended to be younger, and tended more often to be divorced than listed households heads (see also Leuthold and Scheele, 1971). Voluntarily unlisted households "... were less likely to own their homes, and relatively fewer of them resided in suburbia . . ." and fewer had lived at their present address for over two years. Glasser and Metzger (1972; 1975) found that unlisted households (including both voluntary and involuntary) "tend to have younger heads and fewer members 12 years of age or older . . ." and a higher proportion of 18 to 34 years-olds. Leuthold and Scheele (1971) found that unlisted households are much more likely to be black and be city dwellers than listed households in Missouri. Because unlisted households may represent up to 30 percent of all telephone households, the potential bias may be large.

Another problem with the directory as a sample frame is multiple telephone listings within a household. Households with more than one listing have a larger chance of being in the sample. Cooper (1964) found that about 3 percent of all Cincinnati households had more than one listing in the directory. Glasser and Metzger (1972) estimate that 2.2 percent of all households have more than one telephone number and 0.2 percent have three or more numbers.

Alternative Sample Frames

Because of the inadequacies involved in designs using the telephone directory as the sampling frame, researchers have developed alternative sampling methods (Chilton Research Services, 1976; Cooper, 1964; Glasser and Metzger, 1972; Hauck and Cox, 1974; Klecka and Tuchfarber, 1975a; Landon and Banks, 1977; Sudman, 1973). These designs have sought to overcome the problems of unlisted telephones. The alternatives utilize some form of random generation of telephone numbers. This article will evaluate several sampling methods which propose to eliminate the inadequacies of telephone sampling. Researchers must be aware of each sampling method's strengths and weaknesses, to choose the most appropriate one for the research study.

The methods may be divided into two basic classes--designs involving the directory and designs not involving the directory. Sample designs involving the directory begin with a sample of numbers from the directory and modify these numbers to allow the inclusion of unlisted telephone households. Other sample designs, not using the directory at all, combine random digits with some type of sample of existing telephone prefixes.

While both types of methods improve the quality of the sample compared with simple directory designs, they all tend to be less efficient. Because not all randomly

generated numbers are connected with households, the alternative designs develop samples that require more call attempts than the simple directory design. The classic trade-off in sampling between quality and cost is clearly present.

Evaluative Criteria

The sample designs will be evaluated on the following criteria:

1. Bias--the difference between the expected value of the sample statistic and the population statistic. An unbiased sample design is one for which the average of all possible sample statistics of size n would equal the population statistic.
2. Precision--the standard deviation of the sample statistics distribution. Precise sample designs will tend to be closer to the population statistic, for a given sample size.
3. Efficiency--the percentage of all telephone numbers in a sample which do connect with households. Random digit sample designs fabricate telephone numbers, some of which may not be connected with households (unassigned numbers and non-residential numbers). The more non-household numbers in a sample, the less efficient it is. Interview costs are inversely related to sample efficiency.

Of course, the completion rate in a sample of telephone numbers depends on more than the efficiency of the sample. The number and timing of callbacks--to numbers which do not answer or are busy--is critical to the completion rate. With no callbacks, the rate will be lower. Therefore, the efficiency of a sample depends on the sampling method, while the percentage of sample numbers contacted (completion rate) depends on callback procedures as well. The present definition of efficiency is restricted to factors relating only to sample design.

4. Feasibility/Ease of Drawing--sample designs differ on information needed about the population and the effort involved to draw the sample. If some required information is unattainable or the effort is larger than expected benefits, the sample design will not be practical.

Non-Directory Sample Designs

Non-directory random digit samples are basically of three kinds: simple two-stage, two-stage cluster, and stratified. A simple two-stage sample selects a working prefix, including its area code, at random, then adds a four-digit random number. The procedure is repeated until the sample is complete. A two-stage cluster design clusters telephone numbers by prefix (exchange, central office). A sample of prefixes is then selected by either a random or systematic procedure. Within these selected prefixes, four random digits are generated to create the sample of numbers. The stratified sample develops random digits for each working prefix, thereby stratifying the population by prefix. Stratified designs differ on the methods used to allocate the sample to the prefix strata.

While the two-stage designs are commonly used in national studies, the stratification designs have been used in local and regional studies. It is not feasible to stratify by prefix in national samples because there are more than 28,000 working prefixes (Chilton Research Services, 1976). All these sample designs can produce unbiased samples of telephone numbers, assuming that accurate prefix information is available.

Simple Two-Stage Samples

Klecka and Tuchfarber (1975a) and Glasser and Metzger (1972; 1975) have used the simple two-stage design. The only information needed is an enumeration of all working prefixes in the area to be sampled.

To reduce non-response bias researchers have developed methods to deal with not-at-homes and refusals. Klecka and Tuchfarber (1975a) made up to six callbacks of non-answering phones. These calls were made on different shifts to increase the likelihood of contacting a respondent. Glasser and Metzger (1972) made up to 20 callbacks. The great effort in making callbacks permitted the authors to accurately estimate the percentage of working numbers in the entire sample. Glasser and Metzger (1972) calculated that 21.3 percent of all sample numbers were connected with households in 1970. The efficiency of the simple two-stage sample--21 percent--is the lowest for all the sampling methods discussed. The primary cause of the low efficiency is that no allowance is made for non-working banks of numbers within working prefixes. (Each prefix consists of 10 banks of 1,000 numbers each. For example, the 1,000 numbers between 3000-3999 are a bank.) If one were to know which banks of numbers were non-working, the sample could exclude random digits which fall within these banks and efficiency would greatly increase. Unfortunately, information from telephone companies on active banks is not readily available. Some researchers have been able to obtain this information (Brunner and Brunner, 1971; Chilton Research Services, 1965; Cooper, 1964; Klecka and Tuchfarber, 1975a, 1975b) while others have ingeniously used directories to determine working banks (Sudman, 1973).

It is also desirable to reduce the refusal rate in the sample. Klecka and Tuchfarber (1975a) had different interviewers recontact initial refusals. If necessary a supervisor made another attempt to gain cooperation. The authors report that these attempts reduced the number of refusals by 30 percent, to 7.2 percent of eligible households. While these methods are successful in reducing non-response, they may be ethically questionable.

Two-Stage Cluster Samples

Chilton Research Services (1965; 1976) has developed (see also Eastlack and Assael, 1966) a national two-stage cluster sample which increases efficiency over simple cluster samples. A sample of 1,000 prefixes is drawn from a list of all working prefixes. The prefixes are listed according to size within regions of the country. Therefore, a systematic sampling of the 1,000 prefixes yields a regionally balanced sample. From each selected prefix a fixed number (e.g., 50) of four-digit random numbers are produced. Finally, numbers are eliminated which fall within non-working banks, which, contrary to most researchers, Chilton has been able to obtain through local business offices.

Chilton makes three callbacks on no-answer calls and has different interviewers recontact initial refusals. Chilton (1976) estimates that working prefixes are only 22 percent full, as a national average. This fullness is relatively close to the efficiency rating of the simple two-stage samples of 21.3 percent. However, when numbers falling within non-working banks are eliminated, Chilton estimates that 48 percent of their sample numbers connected with telephone households. The efficiency of this technique is about twice as efficient, because non-working banks are eliminated.

The relative precision of these two-stage sample designs is difficult to estimate. The two-stage cluster sample would have some loss of precision due to clustering. Because respondents within a prefix may be more alike

than respondents in different prefixes, the clustering of the sample may produce a larger standard error. However, the balancing of geographic regions and other factors may produce standard error gains over the simple two-stage design. While the Chilton design standard error could be reduced by spreading the sample over more clusters (say, 5,000 clusters of size 10), the gain would probably not offset the added effort of acquiring the non-working bank information.

Stratified Samples

Cooper (1964) and Landon and Banks (1977), using two slightly different approaches, have used stratified telephone samples where every prefix within the population is sampled. These approaches differ on the allocation methods to the prefixes. Both methods require information on working banks, and the second method uses information on the number of telephone households within each prefix. Cooper, after determining all working prefixes and working banks within these prefixes, allocates the sample equally to each prefix bank. Information on working banks might be obtained from the telephone company or could be estimated from the directory. Of course, new banks might be added after the directory is published.

From Cooper's data, the stratified sample using equal allocation to working banks has an efficiency of 41 percent, as measured by eligibles reached. While the efficiency is very close to the two-stage cluster efficiency two factors must be considered: callback procedures and working bank fullness. First, Cooper used no callbacks on no-answers, and one callback on busy signals. Thus, 27 percent of the numbers in the sample resulted in no contact (busy signals and no-answers). Many of these "no contacts" might have resulted in completions if callbacks had been made. Therefore, 41 percent is a conservative estimate of efficiency. Second, Cooper's study was conducted in the greater Cincinnati area. Efficiency scores are influenced by the relative fullness of the working banks. The efficiency of the equal allocation stratified method and the two-stage cluster method, using similar callback procedures, should be equivalent to the average prefix fullness of all working banks. Because Cooper's efficiency with no callbacks was equal to Chilton's, the fullness in Cincinnati was probably higher than the national average.

Landon and Banks (1977) allocated a sample to each prefix, proportional to the number of working phones. This allocation method placed more calls in fuller prefixes and fewer calls in less full prefixes. Therefore, efficiency was increased. However, this information is not always available to the researcher. The efficiency estimate of this allocation method in one study was 48 percent, 2 percent higher than the average fullness of all prefixes, indicating that allocation proportional to size is only slightly better than equal allocation.

The precision of the stratified methods should be better than the two-stage methods because every prefix is sampled. However, if respondents do not differ substantially between prefixes, gains will be small. It must also be pointed out that samples stratified by prefixes are virtually restricted to a local study. It would be difficult to stratify by prefix for a national study. Who wants to work with 28,000 strata?

Because completion probabilities differ between strata and the proportional allocation method unequally divides the sample, the strata means must be weighted to insure that the population estimate is unbiased. In the cited study, the weights did not significantly change the results, probably because the population is homogeneous between strata. However, weighting does slightly decrease precision and does add some work to the analysis. It appears as if proportional allocation, in this

example, gave little increase in efficiency and little loss in precision. The benefits of proportional allocation would be greatest where fullness differences between prefixes are large and differences between strata estimates are small. An illustration of these gains will be presented relative to the following sample design.

Directory Sample Designs

If the geographic area to be sampled is concomitant with the area covered by a telephone directory, then random digit directory sampling may be appropriate. A comparison of the feasibility and ease of drawing directory and non-directory samples will be presented in the next section.

Random digit directory sample designs begin by drawing a sample of numbers from the directory. This sample is often drawn using a systematic procedure. These numbers are then modified to allow all unlisted numbers a chance to be included.

With the systematic sample procedure, an interval is determined by dividing the total number of listings in the directory by the sample size. The interval is employed beginning with a random starting point between one and the interval.

Alternatively, by using the pages in the directory as clusters, and taking the same systematic sample on each page, one can use a hammer-and-nail method. The sample from any page is drawn and a nail is driven through the entire directory at the location of the chosen numbers. In this way, samples are simultaneously drawn from each page. However, the researcher must take care not to nail the directory to the desk (or nail through telephone numbers so they can't be read).

Inverse Sampling With Probabilities Proportional to Size

The inverse sampling method devised by Sudman (1973) systematically picks a set number of prefixes and working banks (his illustration uses 10) from the directory. This procedure gives each listed prefix/working bank combination a probability of inclusion proportional to the number of listed telephone numbers in the directory. The final three digits for each prefix-bank drawn are generated by a series of random numbers. Interviewers are instructed to continue dialing within the bank until a quota (his example was 5) of interviews is completed. Just as in two-stage block sampling all numbers have an equal chance of inclusion.

The efficiency of Sudman's method should be equal to the average fullness of the working banks. However, because numbers are clustered, it is possible to have a large share of the sample from relatively empty (or full) banks. While the theoretical efficiency is equal to average bank fullness, a particular sample using this method may have either a much higher or lower efficiency. Sudman presents an illustrative example reproduced in Table 1. The 10 selected banks required 108 calls to complete 50 interviews, largely because one bank had only 100 working numbers. In this bank alone, 58 calls were placed to complete the quota of five interviews. The expected completion rate was only 44.4 percent (48/108). If Cooper's method had been used, the 108 attempts would have been equally divided among the 10 banks. It can be shown that these calls would have yielded about 82.6 completed interviews for a completion rate of 76.5 percent. The improvement over Sudman's directory sample results from not allocating so many calls to the relatively empty bank. If the 108 calls were allocated proportional to the number of working lines within each bank, the 108 calls would have yielded about 90 interviews for a completion rate of 83.3

TABLE 1

COMPARISON OF SAMPLING METHODS ON NUMBER OF CALLS AND COMPLETION RATES BASED ON SELECTED BANKS

# of Banks	Bank Fullness	P(completion)	Sudman		Cooper		Landon	
			Calls Necessary	Calls* Completed	Calls Necessary	Calls Completed	Calls Necessary	Calls Completed
1	700	.7	6	4.2	10.8	7.56	9.9	6.93
2	800	.8	5	4	10.8	8.64	11.3	9.04
3	800	.8	5	4	10.8	8.64	11.3	9.04
4	850	.85	5	4.25	10.8	9.18	12	10.2
5	850	.85	5	4.25	10.8	9.18	12	10.2
6	850	.85	8	6.8	10.8	9.18	12	10.2
7	900	.9	5	4.5	10.8	9.72	12.7	11.43
8	900	.9	6	5.4	10.8	9.72	12.7	11.43
9	900	.9	5	4.5	10.8	9.72	12.7	11.43
10	100	.1	58	5.8	10.8	1.08	1.4	.14
	7.650		108	48	108	82.62	108	90.04
**Completion Rate				44.4%		76.5%		83.3%

* Calls Completed = P(completion) x Calls Necessary.

** Completion Rate = Calls Completed ÷ Calls Necessary.

percent. The gain is the result of allocating only 1.4 calls to the relatively empty bank and 12.7 calls to the fullest banks. While this illustration is rather extreme--bank fullness ranging from 10 percent to 90 percent--it shows the impact of sampling method and bank fullness on efficiency.

The inverse sampling method is advantageous in that no information is required other than a directory and a random number table. This advantage may make the method much more practical than the non-directory methods. While the inverse method makes no allowances for callbacks, it would not be difficult to include a systematic callback procedure.

Two Random Digits

A simpler procedure for sample selection is outlined by Hauck and Cox (1974). Once the sample of numbers is drawn from the directory, the final two digits are replaced with random digits. This simple design uses no quotas and requires no information from the telephone company. Of course, any new banks or sets of 100 numbers which have been added after the directory is printed cannot be included in the sample.

The efficiency will tend to be slightly higher than previously discussed designs because random digits are used for only the last two numbers. Therefore, non-working sets of 100 numbers are not included. These sets could be included in previous designs when the last three or four digits are randomized. Sample precision should be similar to the two-stage cluster method.

Plus-One Sample

Plus-one samples merely add one to the last digit of each number obtained from the directory. Plus-one has also been called add-a-digit, because researchers can create a new sample by adding another digit to a previously constructed sample.

Because only the last digit is modified, the plus-one method is likely to be most sensitive to omission of new sets of numbers being put in service after the directory is printed. Furthermore, all telephone numbers do not have an equal chance of inclusion. For a telephone number to be included the previous number (e.g., 442-5756 plus one = 442-5757) must appear in the

directory. Sample precision should again be similar to the two-stage cluster method.

To counter-balance potential bias problems, plus-one samples appear to be very effective in avoiding calls to non-working numbers. To test the higher efficiency expectations of plus-one as compared to random digit dialing, two projects conducted in separate communities began with equal numbers of telephone numbers selected either by plus-one or a stratified random digit method, utilizing proportional allocation (Landon and Banks, 1977). The efficiencies (as measured by eligibles reached) of the plus-one subsamples are 18 percent and 11 percent higher than the random digit subsamples. See Table 2. The plus-one design proved to be significantly ($p < .01$) more efficient than the random digit method because of the reduction in the inclusion of non-working numbers. The plus-one subsamples include 16 percent and 9 percent fewer non-working numbers than the random digit subsamples. Because only the last digit is changed the plus-one sample is restricted to working banks and working sets of 100 and even sets of 10 numbers.

TABLE 2

RESULTS FROM TWO DIFFERENT SAMPLING METHODS FOR TWO TELEPHONE SURVEYS

Category	Sterling		Boulder	
	Random Digit	Plus-One	Random Digit	Plus-One
Eligibles	47%	65%	38%	49%
Non-Working	26	10	38	29
Non-Residential	8	8	10	12
Not Connected				
(Busy, No-Answer)	19	17	14	10
	100%	100%	100%	100%
\bar{n}	367	362	454	456
Average Fullness of Banks:		59%		46%

While plus-one appears to be more efficient than random digit dialing, it may potentially be more biased. While

fewer non-working numbers are sampled, the design also may not give all working numbers a chance to be included.

Bias is introduced in plus-one samples to the extent that the response from numbers which have no opportunity to be in the sample would be different from those responses of numbers which could be in the sample. For example, if a group of high income residents began unlisted telephone service at the same time, and if the numbers were assigned sequentially, then all but the first resident would not have a chance of inclusion. If a study were conducted to measure income (or any correlate of income) the study would underestimate the population value. Plus-one may be more efficient than random digit designs. However, there may be idiosyncracies involved in the assignment of phones and the listing of phones which may cause some inherent bias. This bias might be avoided or certainly reduced if the researcher uses a random digit dialing sampling technique. Thus, researchers using plus-one must cautiously apply the design and compare the sample demographics with known population demographics as a check on bias.

Feasibility and Ease of Drawing Directory and Non-Directory Random Digit Samples

Directory and non-directory samples differ on how the sample is drawn and what information is needed to draw the sample. Non-directory samples, depending on type, need to have a list of all working prefixes covered, and to be efficient, a list of all non-working banks. Proportional allocation further requires knowing how many working household numbers are in each prefix. The more efficient the sample must be, the more information is needed. However, the only source of this information is the telephone company. Several researchers have commented on how difficult it is to get information (Cooper, 1964; Glasser and Metzger, 1972; Sudman, 1973). AT&T Long Lines does publish a Distance Dialing Reference Guide which lists all working prefixes in the country by area code. Even though most telephone business offices have the publication, it is only for internal use. The researcher needs special approval to use it. AT&T, Long Lines Division, does make a magnetic tape available with area code and working prefix information. The researcher must contact the local Long Lines Division for the \$43 tape. The request must be approved by the local office and the area vice president before the order is filled. The tapes are updated in January and June, but the researcher must buy each tape separately. Information on non-working banks within prefixes is very difficult to obtain. The only source for the information is the local business office controlling the prefix. Some offices consider the information to be confidential, while others need to seek approval from headquarters. On a local level it may be possible to establish contacts with telephone company officials. These contacts may provide the cooperation necessary to design more efficient samples.

Non-directory samples may be very easily drawn with the aid of computers. For national samples particularly, prefixes and random digits may be generated very quickly and cheaply. Research firms which conduct many studies may find the programming of sample selection to be very time saving. On the other hand, local studies which cannot take advantage of existing programs may find a directory sample quicker and easier to draw. Modified directory samples require no information from any outside source. If the directory is recent newly assigned sets of numbers will be included in the sample. However, because directories are required, national samples are difficult to draw.

It cannot be generally concluded whether directory samples are easier to draw than non-directory samples. The appropriateness of the design depends upon the

availability of necessary information, the scope of the study, and whether the researcher will draw many samples or only one.

Conclusion

Each sampling method described in this paper has the advantage of reducing potential bias compared with simple directory sampling. However, each method is also less efficient and more difficult or infeasible to draw than directory sampling. There tends to be an inverse relationship between efficiency and costs and the reduction of bias. Thus, the researcher must carefully weigh the objectives and budget of a study to determine the optimal balance. This paper has evaluated the relative merits and problems of available sampling methods, so that the researcher can balance the efficiency, costs, and potential bias.

References

- James A. Brunner and G. Allen Brunner, "Are Voluntarily Unlisted Telephone Subscribers Really Different?," Journal of Marketing Research, 8(1971), 121-124.
- Chilton Research Services, "Telecentral Communication: An Innovation in Survey Research," presented to Advertising Research Foundation, 11th Annual Conference, New York, New York, 1965.
- Chilton Research Services, A National Probability Sample of Telephone Households Using Computerized Sampling Techniques (Radnor, Pennsylvania: Chilton Research Services, 1976).
- Sanford L. Cooper, "Random Sampling by Telephone--An Improved Method," Journal of Marketing Research, 1(1964), 45-48.
- J. O. Eastlack, Jr. and Henry Assael, "Better Telephone Surveys Through Centralized Interviewing," Journal of Advertising Research, 6(1966), 2-7.
- Gerald J. Glasser and Gale D. Metzger, "Random Digit Dialing as a Method of Telephone Sampling," Journal of Marketing Research, 9(1972), 59-64.
- Gerald J. Glasser and Gale D. Metzger, "National Estimates of Nonlisted Telephone Households and Their Characteristics," Journal of Marketing Research, 12(1975), 359-361.
- Mathew Hauck and Michael Cox, "Locating a Sample by Random Digit Dialing," Public Opinion Quarterly, 38 (1974), 253-260.
- William R. Klecka and Alfred J. Tuchfarber, Jr., "The Advantages of Random Digit Dialing Surveys: An Empirical Test," unpublished, 1975.
- William R. Klecka and Alfred J. Tuchfarber, Jr., "The Advantages of Telephone Surveys," presented at Annual Meeting of the Midwest Association for Public Opinion Research, Chicago, Illinois, 1975.
- E. Laird Landon, Jr. and Sharon K. Banks, "Relative Efficiency and Bias of Plus-One Telephone Sampling," Journal of Marketing Research, 14(August, 1977).
- David A. Leuthold and Raymond Scheele, "Patterns of Bias in Samples Based on Telephone Directories," Public Opinion Quarterly, 35(1971), 249-257.

Joseph B. Perry, Jr., "A Note on the Use of Telephone Directories as a Sample Source," Public Opinion Quarterly, 32(1968-69), 691-695.

Seymour Sudman, "The Uses of Telephone Directories for Survey Sampling," Journal of Marketing Research, 10 (1973), 204-207.

Alfred J. Tuchfarber, Jr., William R. Klecka, Barbara A. Bardes and Robert W. Oldendick, "Reducing the Cost of Victim Surveys," in Surveys of the Victims of Crime (Ballinger Books, forthcoming).

WHETHER TO AGREE-DISAGREE OR DISAGREE-AGREE:
THE EFFECTS OF ANCHOR ORDER ON ITEM RESPONSE

David Sheluga (student), Jacob Jacoby, and Brenda Major (student)⁽¹⁾
Purdue University

Abstract

Research to date has generally focused on item content rather than item format as a source of questionnaire response bias. A laboratory experiment was conducted to assess the effects of different orderings of verbal and numerical anchors on responses. Results indicate that different rating scale formats differentially influence the degree of item endorsement. An "Agree (=1)---Disagree (=7)" format generates greater degrees of item endorsement than any other verbal and numerical anchor format.

Introduction

A commonly used approach in attitude assessment is to have respondents indicate the extent to which they agree or disagree with one or more attitude statements. Various types of biases may arise when employing such a measurement approach, particularly when a series of items is used and a cumulative score derived. A topic which has received considerable attention is the question of whether such scales are susceptible to response sets or styles, particularly acquiescence (e.g., Cloud and Vaughn, 1970; Elliott, 1961; Jackson, 1967) or "yea-saying" (Couch and Keniston, 1960) response bias.

While debate goes on regarding whether response styles are meaningful or even exist (e.g., Rorer, 1965; Jackson, 1967), those who have attempted to control or adjust for possible agreement response bias have focussed primarily on manipulating item stems through item content (e.g., Jackson, 1967; Elliott, 1961; Wrightsman, 1965; Cloud and Vaughn, 1970). One approach, the method of balanced keying (e.g., O'Neill, 1967; Terborg and Peters, 1974), presents even numbers of positively and negatively worded statements. In contrast to item content, the question of whether response format is sufficient to create, affect, or be used to control agreement response bias has received relatively little attention. As Cook and Campbell (1976, p. 242-3) remarked: "Attitude scales are often presented to respondents without apparent thought... to varying whether the positive end of the response scale appears on the right or on the left of the page."

Considerable evidence exists to show that order effects exist for checklist alternatives (e.g., Belson, 1966; Rugg and Cantrill, 1942). Thus, a reasonable concern is whether such order effects can cause agree or disagree response bias when present in rating scales. Given a seven-point rating scale, would responses differ if the scale ranged from seven (at the top or on the left) to one (at the bottom or on the right), rather than the more common one-to-seven? Similarly, would there be any differential effects if verbal end-point anchors appeared in positive-to-negative (e.g. "agree" to "disagree") rather than negative-to-positive order?

Given that both verbal and numerical anchors are used, another issue meriting consideration is the appropriate

combination of these anchors. To illustrate: is "strongly disagree" more reasonably associated with a "1" or a "7"? In some respects, the issue is analogous to the question of "fittingness" described by Kanungo (1968, 1969).

Accordingly, the present investigation examines the question of whether there are order effects for both verbal and numerical rating scale anchors. Specifically, two studies are described which address the following questions:

1. Do differential, perhaps biasing effects result from having horizontal seven-point graphic rating scales anchored in all four combinations of agree-disagree and 1-to-7?
2. What is the most natural or "fitting" combination of verbal and numerical end-point anchors?

Based upon research cited below, a seven-point scale anchored only at the endpoints was chosen for examination. Considerable literature (e.g., Bendig, 1954; Finn, 1972; Green and Rao, 1970; Guilford, 1954; Lehmann and Hulbert, 1972; Komorita and Graham, 1965; Matell and Jacoby, 1971, 1972; and Symonds, 1924) exists to suggest that 6-7 point scales are optimal for individual level analysis while 3-point scales are adequate for aggregate level analysis. Studies on verbal anchoring suggest that anchors at the end-points are preferred over anchors at every scale point or no anchors at all (Altemeyer, 1970; Bendig, 1953, 1955; Finn, 1972).

Method

Subjects

Subjects were 240 (167 male and 73 female) undergraduates enrolled in Introductory Psychology at Purdue University during the Spring 1976 semester who participated in order to fulfill a class requirement.

Procedure

Measures for the present studies were embedded in a larger set of questionnaires designed to investigate several fundamental issues in questionnaire construction. The combined battery, consisting of eight questionnaires, was administered during one 1-hour session. Each subject responded to all eight questionnaires, with the exception of one questionnaire which always appeared first (described below), the remaining questionnaires were presented in counterbalanced order and were randomly assigned to subjects.

Study 1

An eight-item questionnaire was developed to measure attitudes toward abortion. This topic was chosen based on pre-tests which had indicated that, relative to 12 other topics, abortion was a subject of moderate interest to comparable Purdue undergraduates. Although the eight item stems were identical for all subjects, four different response formats were constructed corresponding to a full crossing of verbal and numerical anchors. The response format for any one subject was the same for all eight items. Each response scale had seven-points. The end-points were labeled both

¹ The authors acknowledge with appreciation the contributions of Amy L. Lubitz and Peggy Whan in planning and executing this study, and Angelo DeNisi, Bernie Dugoni, and John Hollenbach for their statistical assistance.

numerically and verbally as indicated below, while the five intermediate points carried only the intervening numerical descriptors. Subjects were to respond by circling the one number along the scale which best represented their degree of agreement toward that particular attitude statement.

Form	Left Anchor	Right Anchor
1	1 = Agree	7 = Disagree
2	1 = Disagree	7 = Agree
3	7 = Agree	1 = Disagree
4	7 = Disagree	1 = Agree

Form 2 of the questionnaire is provided as Appendix A.

Study 2

This study was directed toward determining whether natural associative tendencies existed between the verbal anchors of Agree -- Disagree and numerical scale end-points of 1 and 7. The design corresponded to a 2 X 2 situation involving the following four scales:

FORM

1a	_____	1	2	3	4	5	6	7	_____
1b	_____	7	6	5	4	3	2	1	_____
2a	Agree	_____	_____	_____	_____	_____	_____	_____	Disagree
2b	Disagree	_____	_____	_____	_____	_____	_____	_____	Agree

This questionnaire was always the first one that each subject received.

Half the subjects received Form 1a while the other half received Form 1b. In both cases, subjects were instructed to write the words "Agree" ("Disagree") and "Disagree" ("Agree") on the lines at the end of the scale (the wording of the instructions was counter-balanced within and across conditions). Form 1b is provided here as Appendix B. Half the subjects in each of these groups then received Form 2a while the other half received Form 2b. Subjects were instructed to write-in scale points from 1 to 7 (or 7 to 1) along the line provided. Again, the wording of the instructions was counterbalanced. Form 2a is provided as Appendix C. Finally, half the subjects engaged in the "assign verbal anchors" task first, while the other half received the "assign numerical anchors" task first.

Results

Study 1

Responses to the eight-item abortion scale were re-coded so that all were aligned in a uniform direction, in terms of numerical and verbal anchors. Responses to the eight items were summed for each subject and a two-way analysis-of-variance for unequal n's was applied. Table 1 presents a summary of these data.

TABLE 1

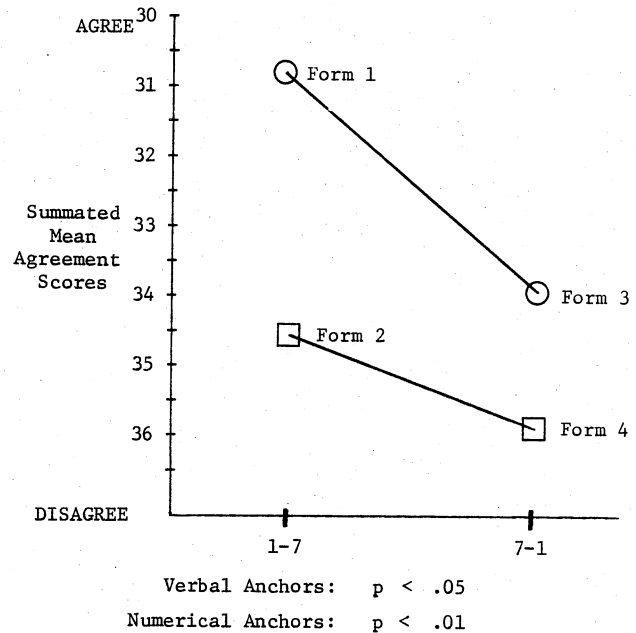
Analysis of Variance Summary for effect of verbal and numerical anchor reversals on total scale scores.

Source	df.	Mean Square	F Ratio	P
Verbal anchors (V)	1	292.27	4.14	< .05
Numerical anchors (N)	1	516.759	7.35	< .01
V X N	1	54.857	.78	N.S.
Error	236	70.276		

Reversals of both the verbal and numerical anchors produced significant differences in the mean scores ($p < .05$ and $p < .01$, respectively), indicating that anchor order does indeed affect the nature of the response. There was no significant interaction effect. Figure 1 depicts these findings.

FIGURE 1

Summary of main effects by two-way ANOVA for reversals of verbal and numerical anchors.



The mean of Form 1 (30.6) differed significantly from the means of Forms 2, 3, and 4 ($\bar{X}_2 = 34.5$; $\bar{X}_3 = 33.7$; $\bar{X}_4 = 35.7$) using a Newman-Keuls range test (Winer, 1970, p. 193).

Data collected using the four formats were next examined to determine whether these formats differed in terms of degree of reliability. Despite the small number of items, a test of split-half reliability produced moderately respectable correlations ranging from .46 to .55 for the four formats. Using Fisher's r-to-z transformation, these differences were found to be insignificant, suggesting that the different forms were equivalently reliable.

Next, responses to each of the eight items were analyzed using separate one-way ANOVAS. Table 2 summarizes these results.

TABLE 2

Analysis of Variance summary for effects of verbal and numerical anchor reversals on item means.

Question	Form	Mean	S.D.	F	p=
#1	#1	3.30	2.01	3.95	.009
	#2	3.95	2.22		
	#3	4.30	1.94		
	#4	4.47	1.94		
#2	#1	6.19	1.34	1.97	.120
	#2	6.00	1.65		
	#3	5.66	1.73		
	#4	6.23	1.06		

TABLE 2 (Continued)

Question	Form	Mean	S.D.	F	p=
#3	#1	4.01	1.94	3.00	.030
	#2	4.69	2.03		
	#3	4.77	2.06		
	#4	5.07	1.95		
#4	#1	2.46	1.74	.69	.560
	#2	2.72	1.70		
	#3	2.27	1.61		
	#4	2.57	1.90		
#5	#1	4.59	2.08	2.08	.100
	#2	5.45	1.77		
	#3	5.03	1.87		
	#4	5.23	2.11		
#6	#1	4.02	2.09	2.93	.030
	#2	5.00	2.18		
	#3	4.64	2.05		
	#4	5.00	2.09		
#7	#1	2.74	1.82	3.09	.030
	#2	3.55	1.98		
	#3	3.66	1.91		
	#4	3.50	1.72		
#8	#1	3.26	2.01	.84	.476
	#2	3.10	1.88		
	#3	3.41	1.88		
	#4	3.65	2.05		

df = 3 in each case

Four of the eight items revealed significant ($p < .05$) differences in means across the four formats, and a fifth item revealed a marginally significant difference ($p = .10$). Examination of individual items revealed that these five statements were relatively "tight," short, and to the point. The three non-significant items, by comparison, had stems which were much longer in word length, and were somewhat more ambiguous.

Study 2

The results of this study are summarized in Table 3.

TABLE 3

Chi-square summary of responses: Study 2

Presented as:	Response Order		Chi Square	P
	AGR-DIS	DIS-AGR		
1-to-7	64	54	.85	N.S.
7-to-1	84	38	17.34	.0001
TOTAL	148	92	13.07	.0001
	1-to-7	7-to-1		
AGR - DIS	68	52	2.13	N.S.
DIS - AGR	84	36	19.20	.0001
TOTAL	152	88	17.07	.0001

Nearly 62% of the subjects felt the most natural response to Form 1 was completing the verbal endpoints in the order of Agree-Disagree ($p < .001$). One-sample Chi-Square analysis showed subjects did not significantly differ in their responses to Form 1a. Associating Disagree-Agree ($n = 54$) to a 1-to-7 scale was almost as frequent as associating Agree-Disagree ($n = 64$) to such a scale. However, responses to Form 1b were significantly different ($p < .001$). That is, responding Agree-Disagree ($n = 84$) to 7-to-1 scale was much more prevalent than responding Disagree-Agree ($n = 38$).

Subjects completed Form 2 by filling-in scale point anchors in the order of 1-7 63.3% of the time ($p < .001$). While responses to Form 2a were not significantly different ($n = 68$ for 1-to-7, and $n = 52$ for 7-to-1), responses to Form 2b (Disagree-Agree presentation format) were significant ($n = 84$ for 1-to-7 vs. $n = 36$ for 7-to-1; $p < .001$).

Discussion

The investigation focused on two related issues. Attention was first paid to the effects on subject responses caused by reversals of verbal and numerical endpoint anchors. Second, given that respondent motivation or willingness of report is a prime condition for successful data collection" (Cannell and Kahn, 1968, p. 537), subject preferences and response tendencies for various scale formats were also assessed. Parenthetically, using subject preference as a criterion may be viewed as reflecting a "consumer (qua respondent) orientation." Within the limitations imposed by the sample, questionnaire, etc., the findings may be summarized as follows:

1. Reversals of either numerical or verbal anchors are sufficient to cause significant differences in response to an Agree-Disagree attitude statement.
2. The Agree-Disagree order resulted in more agree endorsement than did the Disagree-Agree order. Likewise, the 1-to-7 order produced more agree endorsement than did the 7-to-1 order.
3. A rating scale employing an Agree (1) - Disagree (7) order produces significantly more agree endorsement than any other combination of numerical and verbal anchors.
4. Ambiguous or double-barrelled items appeared resistant to being affected by scale reversals.
5. Scale reversals exerted no significant effect on the split-half reliability of the total scale score.
6. Subjects seem to perceive a "natural" association between 1 and Agree, and between 7 and Disagree.
7. Directional tendencies for a scale combining verbal and numerical anchors results in the Agree (1)-to-Disagree (7) format being most frequently selected and considered most "natural."

Closer examination of the data reveals the strength of this preference for the Agree (1)-Disagree (7) format. When the numerical anchors were presented in 1-to-7 order, the subjects were evenly divided as to which verbal anchor order was most appropriate. In other words, the 1-7 order appeared so natural, that either verbal anchor order seemed appropriate when paired with it. This is in contrast to the responses obtained when the presentation format was in the 7-to-1 order. In this condition, subjects overwhelmingly preferred the Agree-Disagree verbal anchor order.

Similarly, when verbal anchors were presented as Agree-Disagree, subjects were evenly divided on the appropriate order for the numerical anchors. The two orders (i.e., either 1-7 or 7-1) seemed equally appealing. In contrast, when the presentation order was Disagree-Agree, the overwhelming choice for appropriate numerical anchor counterpart was the 1-7 order. It is noteworthy that these data are based on a dependent measure requiring more behavioral input from subjects (i.e., "fill-in the end-point as appropriate") than would be required by simple preference ratings of the four formats.

Pre-testing had indicated that, compared to twelve other questionnaire topics, the topic of abortion was of intermediate interest for this student population. This is not unlike topics used in actual survey research. Actual surveys often feature topics which may be of high interest to the sponsor, but are perceived to be bland by the respondents. Thus, the use of a moderate interest topic for this study was considered desirable.

Earlier findings (Elliott, 1961) suggest that there will be a high amount of agree endorsement for low interest topics. Given the predominantly male sample used in this particular study, the topic could be judged as having low personal relevance for these subjects. Again, actual surveys are not often greeted with equal degrees of interest by all respondents. Rather, survey topics are usually differentially relevant to the respondents and therefore vary in interest to the sampling audience.

Our findings indicate that Form 1 facilitates agree endorsement when interacting with a moderately interesting topic having low personal relevance for the subject sample. In part, this may be due to the fact that the ordering of Form 1 was the most "natural" response format, as judged by subjects in Study 2. This familiarity, coupled with moderate or low interest, seems to facilitate the generation of an agree endorsement response style. The findings also suggest that this endorsement can be counterbalanced by reversing either the numerical or verbal anchor order. Usage of either Form 2 or 3 is therefore recommended on this basis. However, both verbal and numerical anchors should not be reversed at the same time, since this seems to cause an inordinate amount of disagree endorsement, as in the case of Form 4.

Use of any of the four combinations does not seem to affect item reliability. This reflects Nunnally's (1970, p. 429) view that format changes will not affect the important psychometric (e.g., reliability) properties of test instruments.

The experimental findings pose a small dilemma. Because respondent frustration may cause response error or "uncooperative" behavior, care should be taken to provide a preferred response format; one with which the respondent feels comfortable. Therefore, Form 1 -- Agree (1)-to-Disagree (7) -- appears most advisable to use. However, caution should be used when employing this format, because it appears that a preferred or "natural feeling" format enhances agree endorsement. It seems advisable that format 1 be avoided when a relatively low interest topic is being investigated. In such instances, Forms 2 or 3 appear more appropriate. Additional research is necessary to more directly investigate the interaction between high and low content interest and the effects of scale reversals.

Appendix A

Attitudes Toward Abortion Survey

Instructions: Please respond to each of the following statements by CIRCLING the number along the line which best expresses your own opinion regarding that statement.

1. It should be left to the woman alone to decide whether or not to have an abortion.

Disagree | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Agree

2. In general, it would solve many social problems if all unwed pregnant women were forced to have abortions.

Disagree | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Agree

3. All states should enforce the Supreme Court decision legalizing abortions.

Disagree | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Agree

4. Pregnant women should be required to obtain authorization from a medical doctor before being permitted to have an abortion.

Disagree | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Agree

5. Abortions should only be legal in the case of a woman who wants an abortion because she was raped.

Disagree | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Agree

6. Abortion is the equivalent of murder.

Disagree | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Agree

7. No abortion should be performed on a fetus over 16 weeks old.

Disagree | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Agree

8. Women on welfare should be able to obtain abortions free of charge through federally financed programs.

Disagree | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Agree

Appendix B

Instructions: Please complete the scale below by writing the words "AGREE" and "DISAGREE" on the dotted line, assigning the words to either end so that the finished scale will be most meaningful and natural for you.

That is, should the number 1 be associated with "Agree" and number 7 with "Disagree"? Or should the number 1 be associated with "Disagree" and the number 7 with "Agree"?

----- . . . | 7 | 6 | 5 | 4 | 3 | 2 | 1 | . . . -----

Appendix C

Instructions: Please complete the scale below by writing the numbers 1 through 7 on the points along the scale, assigning the numbers in the order that will be most meaningful and natural for you.

That is, should the numbers run from 1 to 7, left to right? Or should the numbers run from 7 to 1, left to right?

Agree | | | | | | | Disagree

References

- Altemeyer, Robert A. "Adverbs and Intervals: A Study of Likert Scales," American Psychological Association Proceedings, 1970, 5 (1), 397-98.
- Belson, William A. "On Methods: The Effects of Reversing the Presentation Order of Verbal Rating Scales," Journal of Advertising Research, 1966, 6 (4), 30-7.
- Bendig, A. W. "The Reliability of Self-Ratings as a Function of the Amount of Verbal Anchoring and the Number of Categories on the Scale," Journal of Applied Psychology, 1953, 37, 38-40.
- Bendig, A. W. "Reliability and the Number of Rating Scale Categories," Journal of Applied Psychology, 1954, 38, 38-41.
- Bendig, A. W. "Rated Reliability and the Heterogeneity of the Scale Anchors," Journal of Applied Psychology, 1955, 39, 37-39.
- Cannell, Charles F. and Kahn, Robert L. "Interviewing," in Gardner Lindzey and Elliot Aronson (Eds.) The Handbook of Social Psychology (2nd edition), Vol. 2, Reading, Mass.: Addison-Wesley Publishing Co., 1968.
- Cloud, Johnathan and Vaughan, Graham M. "Using Balanced Scales to Control Acquiescence," Sociometry, 1970, (June) 33 (2), 193-202.
- Cook, Thomas D. and Campbell, Donald T. "The Design and Conduct of Quasi-Experiments and True Experiments in Field Settings," in Marvin D. Dunnette (Ed.) The Handbook of Industrial and Organizational Psychology, Chicago: Rand McNally, 1976, 223-326.
- Couch, Arthur and Keniston, Kenneth. "Yeasayers and Naysayers: Agreeing Response Set As a Personality Variable," Journal of Abnormal and Social Psychology, 1960, 60, 151-174.
- Elliott, Lois L. "Effects of Item Construction and Respondent Aptitude on Response Acquiescence," Educational and Psychological Measurement, 1961, 21(2), 405-415.
- Finn, Robert H. "Effects of Some Variations in Rating Scale Characteristics on the Names and Reliabilities of Ratings," Educational and Psychological Measurement, 1972, (Summer) 32, 255-265.
- Green, Paul E. and Rao, Vithala R. "Rating Scales and Information Recovery -- How Many Scales and Response Alternatives to Use," Journal of Marketing, 1970, 34, 33-39.
- Guilford, Joy P. Psychometric Methods. New York: McGraw-Hill, 1954; Chapter 11, Rating Scales.
- Jackson, Douglas N. "Acquiescence Response Styles: Problems of Identification and Control," in Irwin August Berg (Ed.) Response Set in Personality Assessment, Chicago: Alding, 1967.
- Kanungo, Rabindra N. "Brand Awareness: Effect of Fittingness, Meaningfulness and Product Utility," Journal of Applied Psychology, 1968, 52(4), 290-295.
- Kanungo, Rabindra N. "Brand Awareness: Differential Roles of Fittingness and Meaningfulness of Brand Names," Journal of Applied Psychology, 1969, 53(2), 140-146.
- Komorita, Samuel S. and Graham, William K. "Number of Scale Points and the Reliability of Scales," Educational and Psychological Measurement, 1965, 25, 987-95.
- Lehmann, Donald R. and Hulbert, James. "Are Three-Point Scales Always Good Enough?" Journal of Marketing Research, 1972, 9, 444-46.
- Matell, Michael and Jacoby, Jacob. "Is There an Optimal Number of Alternatives for Likert Scale Items? Study 1: Reliability and Validity," Educational and Psychological Measurement, 1971, 31, 657-74.
- Matell, Michael and Jacoby, Jacob. "Is There an Optimal Number of Alternatives for Likert Scale Items? Effects of Testing Time and Scale Properties," Journal of Applied Psychology, 1972, 56(6), 506-509.
- Munnally, Jum C. Introduction to Psychological Measurement, New York: McGraw-Hill, 1970, Chapter 14.
- O'Neill, Harry W. "Response Style Influence," Public Opinion Quarterly, 1967, 31, 95-102.
- Rorer, Leonard. "The Great Response Style Myth," Psychological Bulletin, 1965, 63, 129-156.
- Rugg, Donald and Cantril, Hadley. "The Wording of Questions in Public Opinion Polls," Journal of Abnormal and Social Psychology, 1942, 37 (October), 469-495.
- Symonds, Percival M. "On the Loss of Reliability in Ratings Due to Coarseness of the Scale," Journal of Experimental Psychology, 1924, 7, 456-61.
- Terborg, James R. and Peters, Larry H. "Some Observations on Wording of Item Stems for Attitude Questionnaires," Psychological Reports, 1974, 35, 463-366.
- Winer, Benjamin J. Statistical Principles in Experimental Design, 2nd edition, New York: McGraw-Hill, 1971.
- Wrightsmann, Lawrence S. Characteristics of Positive Scored and Negative Scored Items from Attitude Scales. Psychological Reports, 1965, 17, 898.

NEAREST NEIGHBOR ANALYSIS: INFERRING BEHAVIORAL
PROCESSES FROM SPATIAL PATTERNS

Sanford L. Crossbart, University of Nebraska
Robert A. Mittelstaedt, University of Nebraska
Gene W. Murdock, University of South Dakota

Abstract

This paper describes a statistic for analysis of spatial patterns generated by behavioral phenomena. Issues associated with use of the method, an application to Whyte's word of mouth study, and other suggested research applications are reviewed.

Problem

The researcher who studies the role of social influence in consumer behavior faces many difficult problems. This is especially true when researching the diffusion or "social contagion" aspect of the innovation adoption process. Using survey methods to isolate the influence of earlier adopters on those who adopt later raises many conceptual and methodological problems, not the least of which are the difficulties in identifying non-verbal influence ("demonstration effects" as Bass, 1969, called them) and the often reactive nature of the measures used (Coughenhour, 1965; Menzel, 1957). However, since the adoption process occurs over time within a social system, many have attempted to infer the presence (or absence) of social influence in observed temporal or spatial patterns of adoption.

It is generally held that an adoption pattern which, over time, conforms to the general shape of the Gompertz logistic function, is evidence of social influence in the adoption process. While the idea is not without its critics, and Midgley (1976) has proposed an alternative model, it is obvious that all studies of temporal patterns require longitudinal data. Some researchers have been able to track an innovation from the moment of its introduction into a social system and have used rather ingenious, nonobtrusive measures of adoption (Coleman, Katz and Menzel, 1957). However, more frequently, researchers have begun their studies only after the innovation has been partially diffused. When historical sales data which allow one to isolate first time purchasers (or assume away repeat purchasers) are available, a pattern can be reconstructed. However, when sales data are not available (as with an innovation which is not an economic good or, possibly, an economic service) or when they contain known but non-identifiable repeat sales (as with goods of relatively short re-purchase cycles) the only course of action is to attempt reconstruction of the process through survey methods.

Rather than examining temporal patterns, one might relate observed spatial patterns of adoption to expectations based on assumptions of social influence. If "contagion" is at work in an adoption process, one would expect to find adopters spatially clustered to some degree. Indeed, this was the argument made by Whyte (1954) in his classic study of air conditioner ownership in a North Philadelphia neighborhood: "As the location of conditioned homes was plotted on a map, a curious distribution pattern began to show up and it could only be explained by the presence of a vast and powerful network (Whyte, 1954, p. 141)".

However, before inferring the existence of social influence from the observation of clusters of adopters, it is necessary to specify the expected spatial distri-

bution in the absence of social influence. Presumably, with neither verbal nor non-verbal interaction among adopters, communications about the innovation would come from an impersonal source (i.e., mass media). Unless there is a predictable relationship between physical location of an adopter within a social system and his/her susceptibility to mass media influence, it would follow that the expected spatial adoption pattern would be that generated by a random process.

Unfortunately, random processes do not produce "uniform" spatial patterns (as Whyte seemed to suggest, p. 141) but would be expected to produce some clusters. Furthermore, as the saturation level of the innovation increased, the incidence of clustering, as a random phenomenon, would also increase. The argument is analogous to that presented by Witt and Bruce (1970) when discussing the use of inter-individual similarities of brand choice as a measure of group conformity. As they correctly argued, any measure of conformity "should reflect not only the degree of brand choice congruence but also the probability that such a group brand choice distribution could have occurred by chance given the market share of the brands involved (p. 533)." To infer social influence from spatial clusters of adopters, the degree of observed clustering should exceed that amount which would be expected by chance alone, at that stage of market saturation.

The purpose of this paper is to describe the nature of a test statistic suitable for this purpose. This statistic, known as "nearest neighbor," will be described in some detail and, by way of example, will be applied to Whyte's North Philadelphia air conditioning data. Rather than dwelling on nearest neighbor calculations or the derivation of associated probability density functions, the discussion will focus on the logic of the analysis and the conceptual and methodological problems which confront researchers who apply the technique. The paper concludes by suggesting some areas of research to which this test statistic might be usefully applied.

The Logic Of Nearest Neighbor Analysis

A nearest neighbor measure is the linear distance between pairs of designated neighboring locations. The analysis can be applied to behavioral phenomena which possess discrete spatial locations which may be mapped as points. Nearest neighbor analysis utilizes the fundamental concept of randomness. A distribution is random when each spatial unit in the area containing the points has an equal opportunity of receiving a point. A nonrandom point pattern is either more clustered than random or more uniform than random. Nearest neighbor analysis reduces the complexities of spatial distributions of observed points to a pattern description called random, more grouped than random, or more clustered than random. Relations between neighboring points are derived under the assumption that such points are randomly distributed in terms of the Poisson distribution. These relations are subsequently used to detect the presence of nonrandomness in given patterns. With the Poisson distribution employed as a standard of comparison, a chi-square comparison has also been employed (Thompson, 1956) but is not utilized in the present

discussion; attention focuses on divergence from randomness along the R scale. R is the ratio of observed average neighbor distance (\bar{F}_a) to expected random neighbor distance (\bar{F}_e) in a spatial distribution

$$R = \bar{F}_a / \bar{F}_e \quad (1)$$

R scale values range from 0 to 2.149. A value of 1.0 provides the standard of randomness, increasing R values are indicative of increasing dispersion leading to a limiting case of regularity, and decreasing R values provide evidence of increasing clustering (Clark and Evans, 1954; Dacey, 1960; Neft, 1966; Taylor, 1977).

Conceptual And Methodological Problems

The nearest neighbor statistic is a useful tool for dealing with certain spatial phenomena. Its successful application depends, of course, on an appreciation of its conceptual and technical limitations as well as the conditions under which its usefulness will be maximized. These considerations are summarized below.

Distinguishing Pattern From Processes

The R scale provides a pattern description. Patterns themselves are the product of underlying processes which develop over time and space. Since patterns provide only static evidence of spacing, they must be approached with notions derived from a rationalization of processes thought to evolve over space. In this sense patterns are no more than abstractions derived by artificially halting dynamic processes. Point patterns are therefore, a synthetic visual expression at a given point in time of processes which continuously operate over space. For this reason any observed point with a spatial distribution represents an event in time and space (Taylor, 1977). Alternative values of R do not necessarily justify the conclusion that either a random or a systematic process is operating. A distribution of observed points provides no direct information about the underlying processes giving rise to the observed distribution (Amedeo and Colledge, 1975). Inferences from patterns to processes can be defended only when processes are considered in forming hypotheses concerning changing R values which are observed over time (Getis, 1964; Pinder and Witherick, 1972; Dawson, 1975).

Limiting Values of the R Scale

As points in space tend to cluster interpoint distances, \bar{F}_a , and hence R, would tend toward zero. The opposite tendency to clustering is dispersion. Taken to the extreme, dispersion results in a limiting regular pattern based on a triangular lattice and an R of 2.149. These theoretical limits to the R scale are rarely observed in practice. The complex processes underlying spatial distributions tend to produce spatial distributions which are more complicated than those reflected by the theoretical limits of the R scale (Pinder and Witherick, 1972). Hence it has been observed (Taylor, 1977, p. 148) that the empirical values of R are likely to range between .33 and 1.67.

Processes Generating Nonrandom and Random Patterns

Extreme cases of clustering and regularity are deterministic in the sense that they have singular R values associated with unique average interpoint distance's which satisfy each extreme case. This determinism does not hold for a randomly produced distribution. A number of random patterns share an R value of 1. In fact, we may purposely use a random procedure to generate a pattern whose R value differs from 1 (Dawson, 1975). In such a case a known random process could be shown to produce a spatial pattern with a tendency toward clustering. This is evidence of a well known phenomena in

inferential statistics, namely that variation in random processes must be taken into account in interpreting stochastic events (Taylor, 1977). The R scale is not merely a descriptive tool which measures a pattern. Various subprocesses are capable of producing R values within the central range of the R scale, resulting in a large number of R values which exhibit small divergence from a random expectation. Attention to this range of random matching (Pinder and Witherick, 1972) forces the researcher to consider the standard error of \bar{F}_e as well as the difference between \bar{F}_a and \bar{F}_e when interpreting a spatial pattern (Clark and Evans, 1954). Still, statistical caution cannot replace an understanding of the dynamics of processes which produce patterns. A narrow view of the application of nearest neighbor analysis (Dacey, 1960) would state that the test is limited to a sensitivity to nonrandomness but cannot be employed to explore hypotheses regarding uniform or concentrated spatial patterns. This view may presuppose the existence of either random variables which produce random patterns or variables which simultaneously contribute to clear cases of regularity or clustering. A broader view (Pinder and Witherick, 1972) maintains that the concept of randomness is not satisfactory as datum. In the real world one might regard a given pattern as a deviation from either or both extremes of regularity and clustering. In the case of diffusion, for example, in certain settings residential sites may bear social meanings which reflect the social leadership character of residents (e.g., larger corner lots). Such a situation could contribute to a hierarchical regular adoption pattern over time. Simultaneously, social contagion would contribute to an observed clustering of adopters. Meanwhile, the psychological effects of mass media could be influencing randomly spaced individual families to make adoption decisions. When such factors operate the random point on a nearest neighbor R scale increases in meaning. The factors which influence the location of adopters are unlikely to operate in a random fashion, but they may distort extreme conditions of regularity or clustering so that the observed pattern is matched by a random distribution. This cannot be attributed to the operation of random forces, but instead might be seen as the significant complex interaction of location and other factors. The role of the researcher is to specify the nature of these factors and their relationships.

Higher Order Analysis

Considering the space or distance between all points and their first or closest neighbor can result in utilizing only a small portion of the information available in a spatial pattern. Imagine pairs of points which are evenly dispersed over space. The distance between any given point and its nearest neighbor would be small. Yet obviously we are not dealing with a clustered pattern as would be indicated by a low R value. Unless each point is considered in relation to all others, the character of the spatial distribution will be obscured by nearest neighbors (Getis, 1964; Neft, 1966; Charlton, 1976; Vincent, 1976). The solution is to identify and average the distance between each point and not only the original nearest neighbor but succeeding neighboring points (second, third . . . nth nearest neighbors). At any order of analysis (so named to refer to the number of succeeding neighbors considered) (Dacy, 1963) observed neighbor distances can be compared to expected distances randomly generated from Poisson distributions. Higher order analysis was made possible by Morrissa's (1954) and Thompson's (1956) derivations of expected distance to nth order neighbors based on calculation of high order probability density functions. In distributions lacking a pronounced spatial bias an order of analysis which is less than n can be employed. Similarly the absence of a pronounced spatial bias supports a direct selection of neighboring points while a bias may support the use of a sectoral decision rule. With this

algorithm the space around each point is divided into the number of sectors equal to the order of analysis and distance from the reference point to the nearest neighbor in each sector is determined (Dacey and Tung, 1963).

Properties of Observed and Expected Distances and the R Scale

The range and interpretation of the R scale depends on the type of analysis employed. The two principal types of more complex nearest neighbor analysis call for either a division of the space surrounding each observed point into k sectors (where $k > 1$) (Taylor, 1977) or measures of spacing between each observed point and its n nearest neighbors (where $n > 1$) (Clark and Evans, 1954; Dacey, 1963). The sectorial method is more sensitive to randomness in a pattern biased toward uniform spacing while raising the order of analysis provides more power efficiency when points are biased toward grouping or clustering (Dacey and Tung, 1962). These types of analysis can be combined, e.g., a space can be divided into k sectors and the distance to n neighbors in each sector can then be determined. For purposes of clarification we shall consider separately the effects of sectorial and n th order nearest neighbor analysis on expected average distances and R scale values. With first order nearest neighbor analysis in which both k and n equal 1, the R scale ranges from values of 0, indicative of maximum clustering through 1.0, indicative of the observed average distances equal to that expected under Poisson conditions ($\bar{F}_a = \bar{F}_e$), to a maximum of 2.149, which represents the case of maximum hexagonal regularity (Neft, 1966; Haworth and Vincent, 1976).

The interpretation of clustering is unaffected by altering the type of analysis in terms of either k or n . This follows from the fact that maximum clustering ($R = \bar{F}_e/\bar{F}_a = 0$) occurs when all points exist at one location and thus the value of observed average interpoint distance would be unaffected by increasing the number of sectors or increasing the number of neighboring distance measures.

By definition, an R value of 1.0 will define a situation in which the observed and expected (under random conditions) average distance are identical. This is obviously unchanged by altering the type of analysis. What will change, however, is the expected average neighbor distances under Poisson conditions. In this sense the distance standard by which the relative randomness of a spatial distribution is judged depends on the type of analysis performed (Morista, 1954; Thompson, 1956; Dacey and Tung, 1962; Dacey, 1963; Taylor, 1977). In a sectorial analysis this outcome is produced by the fact that increasing the number of sectors while holding the order (or n) constant has the effect of increasing the number of nearest neighbor observations. Once the analysis proceeds beyond a single sector ($k = 1$), the expected average distance under Poisson assumptions will naturally increase. Similarly when k is held constant and the n value or order of analysis increases the expected average distance will also increase. Thus while an R value of 1.0 continues to indicate the existence of a random pattern, it masks the important underlying changes in \bar{F}_e which are produced by changes in k and n .

A change in the meaning of R values will occur only in interpreting the regularity of an observed distribution. The R value 2.149 indicates maximized regularity for all values of $n > 1$, just as it does in the first order case where $n = 1$. However, as k or the number of sectors increase an even lattice of points will produce an R value equal to $2.149/k^2$. As k increases the limiting value of R will decrease and actually become less than 1.0. Thus, as the number of sectors around each point is increased the upper limiting case and value of the R scale corresponds, not to that of regularity but, to one

of random expectation under Poisson assumptions. This presents no difficulty if the researcher is aware of scale properties of R under these conditions. This decrease in the R value representing maximum regularity is caused by the fact that, given an even lattice of points, neighbor distances do not increase as k continues to increase and thus the average distance becomes less than that which would be expected under random conditions ($\bar{F}_a < \bar{F}_e$).

Boundary Definition

The parameters for limiting cases of R are based on the assumption of an infinite space or plane. Area is considered only in terms of density and area per point. In reality, of course the boundaries and their effects must be considered. No study area is infinite (Dacey, 1960; Ebdon, 1976).

One associated problem is that of a point whose nearest neighbor lies outside a defined boundary. Ignoring the point results in a pattern description which is biased toward dispersion. Including the point makes the definition of density subject to challenge, but produces a less biased pattern appraisal. For this reason the latter approach is usually taken (Dacey and Tang, 1962; Taylor, 1977). Often, however, the boundary definition is a conceptual rather than technical problem. Defining the area of analysis in terms which correspond to the conceptual nature of the problem being studied will often eliminate the need to consider outlying neighboring points. It is crucial that the spatial context which is under investigation be logically justifiable (Getis, 1964; Bartlett, 1971; Pinder and Witherick, 1972).

This is related to the matter of the size and shape of the study area. For example, an artificially created large study area defined around a small regular pattern of points will provide R values indicating a relatively clustered pattern (Sibley, 1976). In an opposite fashion, shaping the area by drawing boundaries which irregularly follow the observed points and enclosing them in a small space will produce R values indicating regularity. In certain cases, such as that involving linear distributions, special adaptations of nearest neighbor analysis are required (Pinder and Witherick, 1975). In short, the analysis should be modified to meet the peculiarities of the subject and spatial pattern under study. Boundaries must also be defined with attention to the phenomena being studied (Charlton, 1976). Care must be taken to consider how other researchers of similar phenomena have determined relevant boundaries before assertions can be made regarding the comparability of results.

An Application of Nearest Neighbor Analysis

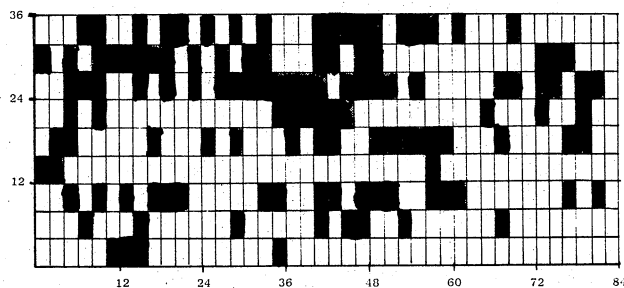
Nature of Data

A classic study by William H. Whyte provides a vehicle to demonstrate an application of nearest neighbor analysis in consumer research. In his *Fortune* article Whyte (1954) described a "powerful communication network" between neighbors as the "web of word of mouth" responsible for the "clusterings" of window air conditioners in a homogeneous Philadelphia neighborhood. These results have been variously referenced in literature dealing with consumer behavior (Engel, Kollat and Blackwell, 1973, p. 404), marketing research (Cox and Enis, 1972, p. 244), diffusion (Rogers and Shoemaker, 1971, pp. 254-55), and marketing principles (Stanton, 1975, p. 105). The findings consist of identified residents of nearly identical \$12,000 row-house type homes, who had adopted window air conditioners. Whyte explains that the neighborhood was physically homogeneous and the occupants (white-collar couples between twenty-five and forty who

earned between \$5,000 and \$7,000) were also fairly homogeneous.

However, the data itself are not presented by Whyte. Results take the form of two aerial photographs of the neighborhood in which the air conditioned homes are marked with X's. One photograph appears in the original article and another photograph, apparently taken at a different time and angle, appears in a later reproduction of the article (Whyte, 1955). Because the data must be gleaned from careful examination of both photos, it has become somewhat altered in later references such as Cox and Enis' research text (1972) which condenses the block lengths, omits homes, and inverts the original distribution. The drawings of blocks used to supplement the photo in Whyte's article represent a purely hypothetical example rather than actual adoptions over time. For illustrative purposes the dominant portion of the neighborhood for which blocks run from left to right within the photographs is transformed in Figure 1 into a spatial distribution of adopters (indicated by darkened areas.) A reference system which defines spatial coordinates is included. The coordinate system was developed from Whyte's photographs by maintaining spacing in proportion to the aerial dimensions of the homes. The center of each air conditioned home provided the spatial coordinates which served as input for the analysis.

FIGURE 1
Whyte's Air Conditioner Ownership Data



Application of Nearest Neighbor Program to Data

A fourth-order nearest neighbor analysis was employed. Given the nature of most distributions, fourth order analysis is sufficient to extract all necessary information from a spatial pattern. To input Whyte's data required that homes be plotted with the streets and "alleys" removed. It was assumed that points would be located in the area within the minimum and maximum coordinates defining the boundaries of the study area, although a researcher has the alternative of specifying boundaries as input parameters. While homes of different adopters cannot possess the same coordinates, this deviation from an assumption regarding a random distribution presents no major problem since any tendency toward clustering or regularity would still be detectable.

Results

Thompson (1956) has shown that in fourth order nearest neighbor analysis the expected neighbor distance is:

$$\bar{r}_e = 1.0937/m^{1/2} \quad (2)$$

and the standard error of \bar{r}_e is:

$$S.E.\bar{r}_e = 0.2774/m^{1/2} \quad (3)$$

where m is the density of observed points, found by dividing N (the number of points) by A (the area in square coordinate units.)

The spatial distribution of air conditioner adopters shown in Figure has the following characteristics:

$N = 123$ points

$A = 2560$ square coordinate units

$m = .04805$ points/coordinate unit

Substituting these values in Equations (2) and (3):

$$\bar{r}_e = 1.0937/((.04805)^{1/2}) = 4.990$$

$$S.E.\bar{r}_e = 0.2774/((.04805)^{1/2}) = 1.265$$

Since the observed mean nearest neighbor distance (\bar{r}_a) is 4.252, the R scale value is computed from Equation (1):

$$R = 4.252/4.990 = 0.8521$$

indicating that the distribution is more clustered than random. To answer the question of whether this is significantly different from the random expectation, it is necessary to recall that the earlier discussion centered on the Poisson distribution as the standard of comparison with the distribution of nearest neighbor distances regarded as approximately normal. Utilizing this line of reasoning, the ninety-five percent limit for the expected mean of N observations would be represented by:

$$\bar{r}_e \pm 1.96 (S.E.\bar{r}_e)/N^{1/2} \quad (4)$$

or

$$\bar{r}_e \pm 1.96 (1.265)/(123)^{1/2}$$

With the limits of expected difference, 4.776 and 5.214, it can be seen that the observed mean nearest neighbor distance (\bar{r}_a) of 4.252 does not fall within the limits of expected distance. (For an alternative derivation of confidence limits under Chi-square assumptions, see Thompson (1956)). Hence it appears that the observed spatial distribution of air conditioner adopters shown in Figure 1 is significantly different from random and exhibits clustering tendencies.

This is not surprising. Our purpose was to apply the R scale to a well-known spatial distribution rather than to challenge Whyte's conclusions regarding the diffusion of air conditioners. In other circumstances involving more complex distributions, neither the spatial pattern nor the outcome would be so obvious.

Research Applications of Nearest Neighbor Analysis

Spatial dimensions of consumer behavior have received insufficient attention. Hopefully this paper will foster increased interest in the consequences of processes which evolve over time and manifest themselves in spatial distributions. It is not intended that the R scale be interpreted as the only or best analytic tool for spatial investigation. Certainly other statistics and theoretical distributions deserve the attention of consumer researchers. For the present the intent has been to highlight the useful nature of nearest neighbor analysis. In addition to applications involving diffusion patterns, the method may be able to shed light on a number of matters including:

- 1) the spatial nature of retail shopping patterns,
- 2) patterns of exploratory movement within retail stores and shopping centers,
- 3) features of cognitive maps representing customer images of the contents of large scale shopping areas,

- 4) the relationship between the perceived locations of retail facilities,
- 5) the distribution of sets of ideal store locations and
- 6) the nature of consumers' social interaction patterns.

These are but a few of the many behavioral phenomena with interesting research possibilities. Application of the nearest neighbor technique will increase our knowledge of its uses and limitations and, as a consequence, may enrich our understanding of consumer behavior.

References

- D. Amedeo and R. G. Colledge, An Introduction To Scientific Reasoning In Geography (New York: John Wiley, 1975).
- M. S. Bartlett, "Two-Dimensional Nearest Neighbor Systems and Their Ecological Applications" in G. P. Patil, E. C. Pielon, and W. E. Waters, eds., Statistical Ecology, Vol. 1 (University Park, Pennsylvania: The Pennsylvania University Press, 1971).
- Frank M. Bass, "A New Product Growth Model for Consumer Durables," Management Science, 15 (Jan., 1969), 215-227.
- B. Charlton, "On Botanists' Boundaries and Buttercups," Area, 8 (1976), 170-1.
- P. J. Clark and F. C. Evans, "Distance to Nearest Neighbor as a Measure of Spatial Relationships in Populations," Ecology, 35 (October, 1954), 445-453.
- J. Coleman, E. Katz and H. Menzel, "The Diffusion of an Innovation Among Physicians," Sociometry, 20 (1957), 253-70.
- C. Milton Coughenhour, "The Problem of Reliability of Adoption Data in Survey Research," Rural Sociology, 30 (1965), 184-203.
- Keith K. Cox and Ben M. Enis, The Marketing Research Process (Pacific Palisades, California: Goodyear Publishing Co., Inc., 1972).
- M. F. Dacey, "A Note on the Derivation of Nearest Neighbor Distances," Journal of Regional Science, 2 (Summer 1960), 81-87.
- M. F. Dacey and T. Tung, "The Identification of Randomness in Point Patterns," Journal of Regional Science, 4 (Summer 1962), 83-96.
- M. F. Dacey, "Order Neighbor Statistics for a Class or Random Patterns in Multidimensional Space," Annals, Association of American Geographers, 53 (1963), 505-15.
- A. H. Dawson, "Are Geographers Indulging In a Landscape Lottery?" Area, 7 (1975), 42-44.
- D. Ebdon, "On the Underestimation Inherent In the Commonly Used Formulae," Area, 8 (1976), 165-69.
- James F. Engel, David T. Kollat and Roger D. Blackwell, Consumer Behavior 2d ed. (New York: Holt, Rinehart and Winston, Inc., 1973).
- A. Getis, "Temporal Land-Use Analysis With the Use of Nearest Neighbor and Quadrant Methods," Annals, Association of American Geographers, 54 (September, 1964), 391-99.
- J. Haworth and P. Vincent, "Maximizing the Nearest-Neighbor Statistic," Area, 8 (1976), 299-302.
- Herbert Menzel, "Public and Private Conformity Under Different Conditions of Acceptance in the Group," Journal of Abnormal and Social Psychology, 55 (1957), 398-402.
- David F. Midgley, "A Simple Mathematical Theory of Innovative Behavior," Journal of Consumer Research, 3 (June, 1976), 31-41.
- M. Morista, "Measuring of the Dispersion of Individuals and Analysis of the Distributional Patterns," Memoirs of the Faculty of Science, Series E, 2 (1954), 254-73.
- D. S. Neft, Statistical Analysis for Areal Distributions, (Philadelphia: Regional Science Institute, 1966).
- D. A. Pinder and M. E. Witherick, "The Principles, Practice and Pitfalls of Nearest Neighbor Analysis," Geography, 57 (1972), 277-8.
- D. A. Pinder and M. E. Witherick, "A Modification of Nearest Neighbor Analysis for Use in Linear Situations," Geography, 60 (January, 1975), 16-23.
- Everett M. Rogers with F. Floyd Shoemaker, Communication of Innovations, (2d ed.) (New York: The Free Press, 1971).
- D. Sibley, "On Pattern and Dispersion," Area, 8 (1976), 163-65.
- William J. Stanton, Fundamentals of Marketing, (4th ed.) (New York: McGraw-Hill, Inc., 1975).
- P. J. Taylor, Quantitative Methods In Geography, (Boston: Houghton Mifflin, 1977).
- H. R. Thompson, "Distribution of Distance to N-th Neighbor in a Population of Randomly Distributed Neighbors," Ecology, 37 (April, 1956), 391-94.
- P. Vincent, "The General Case: How Not to Measure Spatial Point Patterns," Area, 8 (1976), 161-63.
- William H. Whyte, Jr., "The Web of Word of Mouth," Fortune, 50 (November, 1954), 140ff.
- William H. Whyte, Jr., "The Web of Word of Mouth," in Lincoln H. Clark, ed., Consumer Behavior, Vol 2, (New York: New York University Press, 1955).
- Robert E. Witt and Grady D. Bruce, "Purchase Decisions and Group Influence," Journal of Marketing Research, 7 (November, 1970), 533-5.

DURABLE ACCUMULATION: AN EXAMINATION OF PRIORITY PATTERNS

Robert F. Lusch, University of Oklahoma
Edward F. Stafford, Jr., University of Oklahoma
Jack J. Kasulis, University of Oklahoma

Abstract

This paper examines the order of accumulation of discretionary durable goods by consumers. Guttman scalogram analysis was used to identify the underlying priority patterns. A multi-measure procedure was developed to evaluate the reliability and item homogeneity of the derived scale.

Introduction

Over a person's life cycle he or she consumes many goods. Some of these are nondurables, and therefore the consumer derives satisfaction from them over a relatively short span of time. In addition, since these goods are nondurable, the consumer must purchase them with relative frequency. On the other hand, satisfaction from durable goods is derived over an extended period of time and as a consequence these goods are not purchased as frequently.

Most consumers cannot purchase all the nondurable and durable goods they would desire at any given time. This is especially true of durables since they generally involve a larger financial commitment. Therefore consumers will need to decide the order in which they will acquire durable goods. The order of acquisition of durable goods has received little attention by the various parties studying consumer behavior. Instead, the focus has been on how a single good is acquired and on how a consumer selects a brand from among a field of competing brands. Although these two types of investigations have considerable value in the study of consumer behavior, they will not be addressed here. Rather, the purpose of this paper is to explore the order of accumulation of discretionary durable goods by consumers. The order of acquisition of discretionary durable goods is an important area of inquiry since knowledge of acquisition patterns can: (1) aid managers in forecasting sales for consumer durables, (2) provide managers with a framework for directing marketing efforts at consumers at various points in the acquisition process, and (3) provide a framework for more theoretical areas of inquiry such as the adoption of innovations.

Theoretical Background

There are at least three types of models that have been used to explain the acquisition of durable goods. These three models are: (1) growth models; (2) behavioral models; and (3) income consumption models.

Growth Models

Growth models attempt to predict the acquisition of new products by a population. The epidemic growth model which has been adopted from the biological sciences is one such example. This model, which was developed by biometricians to depict the spread of disease, basically postulates a logistics growth curve. When this model is adapted to the consumption of durable goods, the result is a model which can be used to predict the adoption of a new durable good by the population. Briefly, the model suggests that at first only a few people have contact or exposure to a new type of durable good. These few people interact with others and as a result the sales of the product spreads as does an infectious disease spreading

through a closed population. This type of model has been validated for use in predicting the demand for durable goods by Derksen and Rombouts (1937), Roos and Von Szeliski (1939), and DeWoolf (1938). In addition, Bass (1969) has developed a growth model which is a variant on the classic epidemiological model. The model produced fairly good forecasts for the eleven appliance innovations that Bass tested the model with. Several other growth models are discussed by Kotler (1971, pp. 519-560). All of these growth models attempt to depict how the population as a whole adopts new durable goods. Furthermore these growth models depict how a population acquires a single durable good over time. No attention is given to how populations prioritize their acquisition of a set of durables over time.

Behavioral Models

Many constructs and models have been borrowed from sociology and psychology to help understand and predict the purchase patterns of consumers. A large number of these constructs have been combined into comprehensive models of consumer behavior such as the Howard-Sheth model (1969) and the Engel, Kollat and Blackwell model (1968). The major constructs that these models include are: attitudes, motives, personality, perceptual processes, learning processes, and social influences. A complete review of the consumer behavior literature on each of these is obviously unnecessary here. However it is important to point out the lack of evidence on the performance of these models in their entirety. Even more disheartening is the fact that the majority of the available evidence on parts of the models deal with nondurable goods. Consumer behavioralists in general have examined how consumers purchase beer, cigarettes, catsup and cereal, and have deemphasized the study of how consumers purchase durable goods such as houses, air conditioners, and televisions. More important to the discussion at hand, no behavioral research has been directed at how consumers acquire a series of assets (i.e., what are the behavioral determinants of how consumers decide upon which asset to acquire first, second, third, and so forth).

Income Consumption Models

Perhaps the most formal and rigorous theory of consumer behavior available is the classical microeconomic theory of consumption behavior. This theory postulates that given a set of prices for a group of products, consumers will allocate their income to the purchase of these products so as to maximize utility. Importantly, microeconomic theory deals with perishable goods and does not address the problem of how individuals purchase durable goods, such as houses, in which utility is derived not only from consuming the product but from holding stocks of the product. Theil (1954) has suggested that classical economic theory could be reconstructed so that the individual consumer's utility function would depend both on the consumption during a period and stocks held at the end of it.

Derived from the classical economic theory of utility are income consumption curves which allow one to derive the income elasticity of demand. These curves, also frequently referred to as "Engel Curves," have led to the

development of macroeconomic models that quantify the relationship between increases in income and the consumption of goods (durables and nondurables). Predictably, the decision to purchase a durable depends on other factors such as consumer stocks and prices; nonetheless, over the long run there is generally a fairly strong relationship between income and the demand for durable goods. Some examples of the specific relationships for several commodity groups are given in Shubin (1961, pp. 423-441) and Spencer (1968, pp. 146-157). The available evidence suggests that income-consumption models perform better over longer periods of time than for short periods of time and generally do not do well during cyclical periods. In addition, they do a fair job at predicting industry demand for durables but are inadequate for predicting market share within the industry. They do not address the issue of how consumers prioritize their wants for durable goods over time as their income increases.

Order of Acquisition

Three types of models that could be used to explain the acquisition of durable goods have been briefly mentioned. Importantly none of these models has attempted to theorize how consumers go about developing a priority pattern for the acquisition of durable goods. Additionally none have attempted to investigate whether members of a population have similar priority patterns. Obviously, given limited income, consumers will not be able to acquire all the durables they desire at a given moment. Therefore it follows that an order of acquisition will need to exist which will by definition be a temporal concept.

Although classical economic theory does not deal with the consumption of durables, one could reasonably hypothesize that consumers acquire goods over time so as to maximize the present value of their utility function. Thus, given knowledge of future prices, consumer incomes and the utility function, one could theoretically determine the order of acquisition of consumer durables for an individual consumer. Assuming further that all consumers have similar utility structures, one could predict that all consumers would acquire goods in a stated order. For example if there were k durable goods then there would be $k!$ possible patterns of acquisition. Nonetheless, given the above scenario one could predict the order of acquisition; possibly it could be: $D_1, D_2, D_3, \dots, D_k$. Thus, first D_1 would be acquired, then D_2 and so on until D_k is acquired. Initially this assumes that not more than one of a given product could be acquired; but this need not be the case since additional units of a durable good could be treated as separate products. If we wanted to allow for the consumer to purchase up to four of any given durable then given k durables a consumer would have to prioritize $4k$ items.

This notion of the necessity for consumers to prioritize their acquisition of durable goods leads to the hypothesis that populations can be characterized as having a common order of acquisition for many types of consumer durable goods. This implies that consumers will have similar enough utility structures that in general they will acquire consumer durables in the same order.

Several researchers, using a variety of techniques, have examined whether consumers tend to have similar priority of acquisition patterns for sets of consumer durables (Paroush, 1965; McFall, 1969; Hebden and Pickering, 1974). Analysis techniques have included the Guttman coefficient of reproducibility (Paroush, 1965; McFall, 1969), the point correlation matrix (Paroush, 1965), and a matrix of conditional probabilities (Hebden and Pickering, 1974). Generally, it has been found that all three of these methods can be used to construct priority patterns for consumer durables. In addition priority

patterns have been generated for different segments of a market (McFall, 1969; Hebden and Pickering, 1974). The results from this type of analysis suggests that priority or acquisition patterns vary by market segment, however, some overlap tends to exist among the patterns. Finally, priority patterns have been generated based on both intentions to purchase and cross sectional ownership data (McFall, 1969). These two types of analysis tended to generate different patterns of acquisition or priority.

The study at hand will use Guttman scaling, however, importantly multiple measures of scale quality will be computed. Thus although the research reported here is not totally novel, it is unique in that it subjects the Guttman scale to more stringent statistical criteria.

Methodology

The empirical analysis of this paper focuses on the household's acquisition of major discretionary kitchen durables. Included for study are washing machines, dryers, dishwashers, freezers, and microwave ovens. Refrigerators and ranges are the only two major kitchen durables excluded from the analysis. These two kitchen appliances are excluded because they are viewed as necessities for modern household operation, with ownership being inordinately dependent on home ownership. It should further be noted that apartment dwellers may not own washing machines and dryers due to community washers and dryers being provided in the apartment complex. However apartment dwellers were included in the analysis since they were relatively few in number. Furthermore the inclusion of apartment dwellers could not possibly strengthen the scale quality since the presence of community washers and dryers in apartments would induce error into the scale. Thus our procedure is conservative (i.e., by excluding apartment dwellers our measures of scale quality would improve).

Subjects

The data for analysis was obtained from the DRP/OPUBCO Continuing Consumer Audit. The Distribution Research Program (DRP) at the University of Oklahoma and the Oklahoma Publishing Company (OPUBCO) collaborate in the collection of data on the purchasing behavior of individuals in the Oklahoma City SMSA.¹ Reported in this paper are some findings of the 1975 Audit which includes 1854 respondents from a stratified random sample. The sample is representative of the populations in geographic regions in the OKC SMSA which includes subjects from urban, suburban, and rural areas. It includes all types of dwelling units--houses, apartments, condominiums, trailers, etc. New samples are drawn each year with the distribution of the sample reflecting population changes in the strata.

Procedure

The Consumer Audit questionnaire is administered in a personal interview with both the male and female heads of the household responding. Each visit is a "cold" call with no pre-visit contact made with the potential respondent to request cooperation; i.e., the first contact is when the doorbell is rung. Not-at-home families are revisited at different hours of the day four times before a substitute respondent is designated for the interview. The data is collected continuously throughout the year by professional interviewers under close supervision. Each week completed questionnaires are returned

¹The authors wish to thank Phil Stout and Jim Williams of the Oklahoma Publishing Company for their continued cooperation in the collection and use of the Continuing Consumer Audit data base.

for processing. Telephone callbacks are made within three days of the return to verify the data collection. In addition, subjects may be telephoned again to obtain clarification of responses if need be.

Analysis

The stock of durables that a household possesses can be characterized by a multivariate distribution of 0's and 1's. A value of "one" would depict possession of the durable. A value of "zero" would represent non-ownership. If it is true, as hypothesized, that consumers have relatively similar utility structures for discretionary kitchen durables, the data on durable consumption behavior can be described in terms of a unidimensional "scaling" model.

The approach used in this paper is similar to that of Paroush (1965). In dealing with five household durables, a logically consistent conceptual model would be theoretically characterized by the pattern exhibited in Table 1. This indicates the order of acquisition to be D_1 , D_2 , D_3 , D_4 , D_5 . If it is observed that each consumer fits into any one of these patterns (rows), then we could transform the multivariate data into a unidimensional scale. Thus, by only knowing the last durable acquired, one can perfectly predict a consumer's total stock of durables. In other words, if the last durable added to one's stock of durables is D_4 , then we would know that this consumer also possesses D_1 , D_2 , D_3 and not D_5 .

Table 1
Theoretically Perfect Patterns of Durable Ownership

Subject	Durables				
	D_5	D_4	D_3	D_2	D_1
S_5	1	1	1	1	1
S_4	0	1	1	1	1
S_3	0	0	1	1	1
S_2	0	0	0	1	1
S_1	0	0	0	0	1
S_0	0	0	0	0	0

Now that the theoretically pure or ideal situation has been developed, it should be clear that not all consumers will acquire a set of durables in the same pattern. For whatever the reasons, the perfect durable acquisition scale will not be able to characterize all people since deviants will inevitably exist. The task is to determine whether divergence from the ideal model is a function of relatively unimportant aberrations or whether the divergence is sufficiently large for the perfect model to be considered unrealistic for the real world. Thus, some measure of scalability is needed to empirically test the appropriateness of the model.

Guttman (1971) developed a scaling model called scalogram analysis which may be applied to this task even though it was originally devised for a different purpose (attitude measurement). In the discussion of attitudinal dispositions, and the role of scalogram analysis, Guttman has stated that "...the universe is said to be scalable for the population if it is possible to rank the people from high to low in such a fashion that from a person's rank alone we can reproduce his response to each of the items in a simple fashion." (Guttman, 1971, p. 188).

From the Guttman perspective, an attitude scale should possess two properties: (1) unidimensionality and (2) cumulativity. A unidimensional scale is one in which

the component items measure movement toward or away from a single underlying object. A cumulative scale is one in which the components can be ordered by degree of difficulty; i.e., if a respondent replies positively to a more difficult component of the scale, then he would always respond positively to each of the less difficult scale components. For example, consider the following two simple attitudinal statements: (1) I would use Brand X coffee if it were given to me as a sample. (2) I would buy Brand X coffee in the supermarket if I could. A positive response to both statements would indicate a stronger disposition toward Brand X than a positive response to only one of the statements (unidimensionality). A positive response to Statement 2 should always result in a positive response to Statement 1, but not necessarily vice versa (cumulativity).

Guttman scaling has been traditionally used with cross sectional data in order to rank an individual's attitude toward an object. In the case at hand, it is desired to use Guttman scaling to model the temporal phenomenon of consumers acquiring discretionary durable goods. Although the data used in this study is cross sectional, it is possible to scale the underlying temporal phenomenon. Since the sample represents a true cross section of the entire population of the Oklahoma City SMSA, individuals at all stages in the order of acquisition process will be present in the sample. Thus, because this cross section of individuals is at various stages in the acquisition process, conclusions can be drawn about the order of acquisition over time.

From the above discussion, the appropriateness of scalogram analysis in examining our hypothesis is evident. The thesis of this paper is that the five discretionary kitchen durables mentioned earlier tend to be acquired in a designated priority pattern with the "more difficult" durables being acquired only after the "less difficult" appliances. In this context, a lesser degree of difficulty is synonymous with higher levels of expected utility derived from the ownership of the appliances. Among the various scaling techniques available, the Guttman approach is almost unique in its possession of this cumulative property (Nie, et al., p. 529).

The various techniques used to analyze the data collected for this study are detailed in the Appendix to this paper. These techniques include the original measures suggested by Guttman (1971); the Kuder-Richardson Equation 20 (Kuder and Richardson, 1937); the approach of Loevinger (1948); the extensions of the Kuder-Richardson and Loevinger works, suggested by Horst (1953); and the unique approach suggested by Green (1956). (A Technical Appendix, detailing the calculations performed using these techniques in order to derive the statistics reported in Table 3, is available on request from the authors.) The results of applying these techniques are presented in the next section of this paper.

Results and Analysis

As mentioned earlier, for the five discretionary kitchen durables, there are 32 ownership situations, yet only one pattern of acquisition represents the perfect scale. (There are $k+1=6$ ownership situations represented by this perfect scale.) To examine the extent of deviation from the perfect scale pattern, a variety of techniques (described in the Appendix) were used. A statistical summary of the data used to generate the Guttman scalogram, and to generate the scale reliability measures associated with these analytical techniques, is presented in Table 2.

A brief interpretation of Table 2 is in order. The scale rankings (row labels) are determined, for each respondent, according to the Guttman "perfect scale" technique. That is, a respondent receives a scale score equal to the

Table 2
 Statistics Summary of the Guttman Scale Analysis for Consumer
 Acquisitions of Discretionary Kitchen Durables^a

Scale Ranking	Durable Good										Totals		
	Microwave		Oven		Freezer		Dishwasher		Dryer			Washer	
	0 ^b	1 ^b	0	1	0	1	0	1	0	1		0	1
5 ^c	0	24	0	24	0	24	0	24	0	24	0	24	24
4	322	22	20	324	2	324	0	344	0	344	0	344	344
3	476	8	308	176	168	316	16	468	0	484	0	484	484
2	413	1	365	49	400	14	56	358	8	406	8	406	414
1	181	0	144	37	163	18	176	5	60	121	60	121	181
0	346	0	346	0	346	0	346	0	346	0	346	0	346
Sums	1738	55	1183	610	1079	714	594	1199	414	1379	1793 ^d		
Errors	0	31	20	262	170	32	72	5	68	0			660

^aBasic Guttman analysis was obtained using SPSS(Nie, et al., 1975, pp. 528-539).

^bA 0 column gives the number of respondents not owning the durable;
 A 1 column gives the number of respondents owning the durable.

^cScale rankings are assigned according to the Guttman "perfect scale" technique.

^dOf the 1854 respondents surveyed, 61 failed to complete one or more questions related to the five items in this scale, and thus were discarded from the sample.

number of positive responses (1's) he gives to the items on the scale. For this study, a ranking of 3 indicates that the respondent possessed three of the five kitchen durables. Each column of Table 2 actually consists of two columns of numbers; the left-hand sub-column represents the number of respondents who did not possess the item represented by the overall column, for a given scale ranking, while the right-hand sub-column represents the number of respondents who did possess that item, for the given ranking. For example, of the 344 respondents who received a scale ranking of 4, two did not possess a dishwasher, while 342 did. Note that the sum of the sub-column entries, for any column, equals the total number of respondents in this study. Further, the sum of the sub-column elements for any row equals the number of respondents receiving that row's scale ranking.

The "easiest" items (most possessed) start at the right of Table 2, and progress to the "most difficult" items (least possessed) at the left of this table. Thus the "ideal" order of acquisition of kitchen durables would be: washer, dryer, dishwasher, freezer, and finally microwave oven. The values in the box of each sub-column of Table 2 represent the number of respondents that deviated from the perfect scale, for the item represented by the overall column. For example, of the 181 respondents scaled as a '1', 60 did not own the first item in the scale, a washer. Of these 60, five owned a dryer as their single durable; 18, a dishwasher, and 37, a freezer. These item errors occurred either by respondents possessing an item of greater difficulty than their scale value suggests, or by not possessing an item that would normally be owned, given their assigned scale value. The various row and column totals from Table 2 are used to generate the reliability measures, presented in Table 3, and discussed below.

Examination of the performance measure values presented in Table 3 provides strong support for the hypothesis of this paper that consumers tend to acquire discretionary kitchen durables in a pattern that can be detected through scalogram analysis. The Guttman measure of scale reproducibility exceeds 0.85, the minimum level suggested by Guttman for claiming a valid scale. Further, Green's

index of consistency substantially exceeds his suggested minimum value of 0.5; and the coefficient of scalability is well above 0.6. These results suggest that the proposed scale, based on the hypothesis stated earlier, is both reproducible and consistent.

Table 3
 Reliability Measures of the Discretionary
 Kitchen Durables Scale

Analytical Technique	Total Sample	Split Half I	Split Half II
Guttman			
Coefficient of Reproducibility	0.926	0.933	0.920
Minimum Marginal Reproducibility	0.734	0.737	0.731
Percent Improvement	0.193	0.196	0.189
Coefficient of Scalability	0.724	0.745	0.703
Loevinger			
Index of Homogeneity	0.659	0.675	0.644
Kuder-Richardson			
Standard KR ₂₀	0.688	0.689	0.688
Corrected KR ₂₀	0.850	0.854	0.840
Green-B			
Index of Reproducibility	0.963	0.966	0.960
Expected Coefficient of Reproducibility	0.895	0.898	0.893
Index of Consistency	0.646	0.668	0.624
Sample Size	1793	893	900

Drawing conclusions from the Loevinger Index, the Green Index, or the corrected Kuder-Richardson Equation 20 measure is considerably more difficult. To date no suggestions of reasonable minimum levels for any of these performance measures has been found anywhere in the literature. Despite this apparent lack of historical/theoretical guidance, the procedure described below for combining the Guttman, Green, Loevinger, and Horst measures for assessing the validity and reliability of scales developed from a set of dichotomous variables has been chosen. Future research efforts need to identify reasonable bounds of the Loevinger, Green, and Horst

measures.

First since both the Guttman and the Green Reproducibility Coefficients measure the same phenomenon, both should approach, if not exceed 0.85. Second, the Coefficient of Scalability should exceed 0.6, and/or Green's Index of Consistency should exceed 0.5. Further because the Loevinger measure and the Green Index of Consistency both attempt to evaluate the homogeneity of a scale, the Loevinger Index should also exceed 0.5 as a minimum. The value of the Horst measure, then, will also exceed 0.5 because it is always a multiple (≥ 1.0) of Loevinger's Index. If all of these conditions are met, it is concluded that a valid, reliable scale has been identified.

For this current study, all sample values of the various reliability and homogeneity measures exceeded the stated minimums. Hence it was concluded that a reliable and valid scale has been identified for the acquisition of discretionary kitchen durables.

To further test the internal validity of the scaling procedures reported above, the sample of 1854 respondents was split into two halves. The respondents were split according to the odd or even last digit of their sequential identification number to provide two subsamples. Since the completed questionnaires, gathered over many months of 1975, were numbered as they were collected from the interviewers, neither split-half sample was biased by seasonal factors. The reliability statistics calculated for each split-half are reported in Table 3. Inspection of these statistics strongly suggests that the sampling procedures used in this study were indeed random, unbiased and internally valid.

For the present, Paroush's notion of examining an item correlation matrix for simplex structure has been excluded from this scale evaluation procedure. This was done because it is not clear just which of the many available procedures for calculating correlations should be used. The method used by Paroush provides a significantly different matrix than the Yule's procedure employed in SPSS; nonetheless, each method produced a matrix that had a nearly perfect simplex structure. The Paroush version of the point-correlation matrix for the items examined in this study is presented in Table 4. The only flaws in this matrix are both associated with the item "freezer." This same phenomenon also appears in the data shown in Table 1.

Table 4
Point-Correlation Matrix for the Discretionary
Kitchen Durables Scale Items

Item	Washer	Dryer	Dishwasher	Freezer	Microwave Oven
Washer	1.000	.750	.389	.271	.090
Dryer		1.000	.459	.260	.118
Dishwasher			1.000	.291	.146
Freezer				1.000	.057
Microwave Oven					1.000

Discussion and Conclusions

This paper expanded earlier work on an intuitively appealing application of Guttman scaling. Is the application of any value; and if this application has value, is its value theoretical, managerial or both? We believe that the ability to characterize the population's multivariate acquisition of goods by a single scale offers both immediate managerial and theoretical possibilities.

Managerially, it can be said that the ability to scale

the order of acquisition of a set of durables allows managers a convenient way to estimate market potential, or forecast sales. For example, the potential purchasers of microwave ovens are those consumers that have already acquired a washer, dryer, dishwasher, and freezer. Those consumers that have only purchased a washer are not presently potential customers for microwaves until they first purchase a dryer, dishwasher and freezer. Obviously, therefore, the market potential changes as the population progresses through the order of acquisition process. In regard to forecasting sales, Brown, Buck and Pyatt (1965) have developed a technique utilizing priority patterns that can be used to improve the sales forecasts for consumer durables.

Second, it would be possible for managers to use the order of acquisition phenomenon to direct marketing efforts. In short, if microwaves were being sold, then marketing dollars should not be wasted on consumers who have only purchased a washing machine. In this case, direct mail could be focused on recent freezer purchasers, inasmuch as their next most likely new discretionary durable would be a microwave oven.

Turning to the more theoretical applications, a researcher could use the order of acquisition scale to investigate its potential for studying the adoption of innovations. Perhaps a person's scale score could be a determinant of him adopting a new innovation. For example if one examined those consumers that have adopted trash compactors, one would possibly find that the adopters were the individuals that had a high scale score (i.e., had already acquired all other discretionary kitchen durables).

Also on a theoretical plane, one could investigate the concurrent and/or predictive validity of the scale. The scale could be investigated to see how it correlates with, and its ability to predict, other behavioral constructs.

Finally, theoretical research could be directed at techniques for minimizing the error in the scale. This may be possible by segmenting the population to see if order of acquisition varies by segments. For example, individuals entering adulthood may acquire durables in several different orders depending on such things as marital status and social class. By constructing separate Guttman scales for each segment more perfect scales should result.

In summary, it can be concluded that the study of the order of acquisition of durable goods offers a promising area to all the respective parties interested in the study of consumer behavior. Coupled with this, it appears as though the Guttman scale can be a fruitful technique for the study of the order of acquisition of durables.

Appendix

This Appendix presents a discussion of the various techniques available for analyzing the data gathered in conjunction with the identification of a unidimensional scale. Except for the Guttman approach, which is presented first, these techniques are presented in the order in which they appeared in the literature.

Guttman

The most frequent analytical technique used in scalogram analysis and the most common one operationalized in the more popular computer statistical packages is the Coefficient of Reproducibility (C_R). Related measures are the Minimum Marginal Reproducibility (C_{MMR}), Percent Improvement ($\%I$), and Coefficient of Scalability (C_S). Guttman's C_R indicates how well the data approximates a

perfect scale by examining the extent to which positive replies to "more difficult" scale items are associated with the positive replies to those which are "less difficult". Mathematically, this statistic is as follows:

$$C_R = 1.0 - \frac{\text{total number of errors}}{\text{total number of responses}}; 0 \leq C_R \leq 1.0 \quad (1)$$

Errors are said to occur when a positive response is given to a "more difficult" item and not to a "less difficult" one. As a rule of thumb, Guttman assumes a valid scale if C_R exceeds 0.85 (Guttman, 1971, p. 176).

Subsequent research noted that C_R is biased upward to the extent that the dichotomized items have extreme distributions. The C_{MMR} measure was therefore devised to indicate the minimum coefficient of reproducibility that could occur for a scale given the proportion of respondents replying positively to each of the items. It is calculated as follows:

$$C_{MMR} = \frac{\sum \text{Maximum marginal for item } i}{\text{total number of responses}} \quad (2)$$

It should be noted that C_{MMR} is in itself, a meaningless statistic. Its value comes from an examination of its relationship to C_R . This relationship is examined through the %I statistic. It indicates the extent to which C_R is due to response patterns rather than to the underlying associations of the scale items. In this regard, the percent improvement is akin to the error or measurement variance of variance analyses techniques. Mathematically it is as follows, with the larger the value of %I, the stronger the indication of a valid scale.

$$\%I = C_R - C_{MMR} \quad (3)$$

The final Guttman statistical ratio is C_S . It is calculated as follows:

$$C_S = \frac{\%I}{1.0 - C_{MMR}}; 0 \leq C_S \leq 1.0 \quad (4)$$

The denominator of this formula indicates the largest amount of percent improvement which can be obtained and the numerator represents the actual amount obtained. Therefore C_S is a measure of how close the scale has come towards achieving the maximum improvement possible. Values above 0.6 are considered necessary for scale validity (Nie, et al., 1975, p. 533).

Kuder-Richardson

Several years prior to Guttman's original reports on his scalogram procedures, Kuder and Richardson (1937) described a method of estimating the reliability of tests and scales. They described several special cases, one of which has proven useful for testing for this current study. As with their other cases, the Kuder-Richardson Equation 20 (KR_{20}) is based on the variances of the individual items of a test or scale, and on the total variance of the entire scale. This equation was given as:

$$KR_{20} = \frac{k}{k-1} (1 - V_I/V_T); 0 \leq KR_{20} \leq 1.0 \quad (5)$$

where: k = the number of items in the scale
 V_I = the sum of the variances associated with each individual item of the scale
 V_T = the total variance associated with the entire scale

Loevinger

Subsequent research has noted two defects in the KR_{20} approach. Loevinger (1948) argued that scale reliability formulas, like KR_{20} , estimate item homogeneity as well as scale or test reliability. Furthermore, Loevinger (1947) pointed out that KR_{20} has an upper limit of 1.0 only when all items of the scale were of equal difficulty. As a consequence, Loevinger developed an alternate measure of test reliability called the Coefficient of Homogeneity (L_I). L_I produces identical results as KR_{20} when all items are of equal difficulty, but corrects for situations where there is a dispersion of item difficulty. A computationally useable formula for L_I is as follows (Horst, 1953):

$$L_I = \frac{V_T - V_I}{V_M - V_I}; 0 \leq L_I \leq 1.0 \quad (6)$$

where: V_M = maximum possible variance given the distribution of item difficulty
 V_T, V_I = same as for KR_{20}

Horst

Later discussion by Horst (1953) demonstrated that L_I really estimated the average item correlation for a test or scale, corrected for dispersion of item difficulty. Additional sophistication by Horst resulted in a better measure of reliability in that it incorporated total scale reliability. His Corrected Kuder-Richardson Equation 20 (KR^*_{20}) is calculated as follows:

$$KR^*_{20} = \frac{V_T - V_I}{V_M - V_I} (V_M/V_T); 0 \leq KR^*_{20} \leq 1.0 \quad (7)$$

Green

A different approach to estimating the reliability of a scale was introduced by Green (1956). It depends not on the calculations of item and test variances, but rather on the summary statistics of the sampling results obtained when a scale is designated. A researcher first estimates a scale for the items of interest, ranking them in order of ascending difficulty with item 1 being the least difficult. Next, the number of positive and negative responses are identified for each item. Although there are $k-1$ orders of possible error responses for a scale of k items, Green argued that, at most, just the first two orders contributed significantly to the reliability coefficients of a scale. A first-order error would be one where a respondent answered positively on an item of greater difficulty, say item $g+1$, while he responded unfavorably on item g . The index g is calculated over the ranked set of items. A count of all such first-order errors for the g th item would be denoted as $n_{g+1,g}$, meaning this many respondents possessed the $g+1$ st item, but not the g th item. Extension of this discussion and notation to second-order error counts is obvious.

Green presented two alternative estimates of the reliability of a scale which he called Rep_A and Rep_B . He showed the equivalence of these measures with the assumption that item error counts are independent statistics. Because it is computationally much less difficult, the Rep_B estimate has become more popular among social scientists. The Green-B Index of Reproducibility (R_{GB}) is as follows:

$$R_{GB} = 1.0 - \frac{1}{Nk} \sum_{g=1}^{k-1} n_{g+1,g} - \frac{1}{2Nk} \sum_{g=2}^{k-2} (n_{g+2,g})(n_{g+1,g-1}) \quad (8)$$

where: N = the number of respondents to the test (scale)

k = the number of items in the test (scale)
 n = the number of errors for the particular item-order interaction
 g = the item index across the set of ranked items for the scale

Although not explicitly stating so, Green implied that this index should be tested as though it were the original Guttman R_G . Thus, a value of R_{GB} in excess of 0.85 suggests that a valid scale has been identified.

Green also introduced the Index of Consistency (I) to serve as a measure of the homogeneity of the developed scale. This index is calculated as follows:

$$I = \frac{R_{GB} - R_I}{1.0 - R_I} \quad (9)$$

where: R_{GB} = Green B Index of Reproducibility (defined above)
 R_I = Expected Coefficient of Reproducibility (defined below)

R_I is the scale reproducibility that is expected by chance if a set of items possessed their observed popularities, and they are also mutually independent. Thus, this index indicates the extent to which an estimated R_{GB} is representative of a valid scale for the observed items. A negative value for I indicates a degree of negative correlation among some or all of the items in the scale. The equation for R_I is as follows:

$$R_I = 1.0 - \frac{1}{N^2} \sum_{k=1}^{k-1} (n_{g+1})(n_g) - \frac{1}{4} \sum_{k=2}^{k-2} (n_{g+2})(n_{g+1})(n_{g-1}) \quad (10)$$

where: N, k, n, g = same as for R_{GB}

Paroush

In addition to the approaches just discussed, another method for judging the validity of a scale has been suggested by Paroush (1965). In this approach, the scale items are also ordered by their ranked popularity. A point-correlation matrix for each pair of ranked items is then calculated. If this matrix has a simplex structure, the hypothesized scale is accepted. A simplex structure is one where the values of the matrix entries decrease as one moves away from the matrix diagonal, either horizontally or vertically.

Summary

Discussed have been several measures of scale reliability or reproducibility found in the literature of the social sciences. These measures vary not only in how they are constructed, but also in their ability to deal with scales where items are believed to differ in degree of difficulty. It is not the intent of this Appendix to discuss in detail the relative merits of these measures; rather, each of them has been calculated for this current study as a measure of scalability of discretionary kitchen durables.

References

Frank M. Bass, "A New Product Growth Model For Consumer Durables," Management Science (January 1969), 215-227.
 D.A. Brown, S.F. Buck and F.G. Pyatt, "Improving the Sales Forecast for Consumer Durables," Journal of Market-

ing Research, 2 (August 1965), 229-234.

J.G.D. Derksen and A. Rombouts, "The Demand for Bicycles in the Netherlands," Econometrica, 5 (1937), 295.

P. DeWoolf, "The Demand for Passenger Cars in the United States," Econometrica, 6 (1938), 113.

James F. Engel, David T. Kollat and Roger D. Blackwell, Consumer Behavior (New York: Holt, Rinehart and Winston, Inc., 1968).

Bert F. Green, "A Method of Scalogram Analysis Using Summary Statistics," Psychometrica, 21 (March 1956), 79-88.

Louis Guttman, "The Cornell Technique for Scale and Intensity Analysis," in Gene F. Summers, ed., Attitude Measurement (Chicago: Rand McNally & Company, 1970), 187-202.

Louis Guttman, "A Basis for Scaling Qualitative Data," in Gene F. Summers, ed., Attitude Measurement (Chicago: Rand McNally & Company, 1970), 174-186.

J.J. Hebden and J.F. Pickering, "Patterns of Acquisition of Consumer Durables," Oxford Bulletin of Economics and Statistics, 36 (May 1974), 67-94.

Paul Horst, "Correcting the Kuder-Richardson Reliability for Dispersion of Item Difficulties," The Psychological Bulletin, 50 (1953), 371-374.

John A. Howard and Jagdish N. Sheth, The Theory of Buyer Behavior (New York: John Wiley & Sons, Inc., 1969).

G.F. Kuder and M.W. Richardson, "The Theory of the Estimation of Test Reliability," Psychometrica, 2 (September 1937), 151-160.

Philip Kotler, Marketing Decision Making A Model Building Approach (New York: Holt, Rinehart and Winston, 1971).

Jane Loevinger, "A Systematic Approach to the Construction and Evaluation of Tests of Ability," Psychological Monographs, 61 (1947), no. 285.

Jane Loevinger, "The Technique of Homogeneous Tests," The Psychological Bulletin, 45 (1948), 507-529.

John McFall, "Priority Patterns and Consumer Behavior," Journal of Marketing, 33 (October 1969), 50-55.

Norman H. Nie, C. Hadlac Hull, Jean G. Jenkins, Karin Steinbrenner and Dale H. Bent, Statistical Package for the Social Sciences (New York: McGraw-Hill, 1975).

Jacob Paroush, "The Order of Acquisition of Consumer Durables," Econometrica, 33 (January 1965), 225-235.

C.F. Roos and V. Von Szeliski, "Factors Governing Changes in Domestic Automobile Demand," The Dynamics of Automobile Demand (New York: General Motors, 1939), 21.

John A. Shubin, Managerial and Industrial Economics (New York: Ronald Press Company, 1961).

Milton H. Spencer, Managerial Economics (Homewood, IL: Richard D. Irwin, Inc., 1968).

H. Theil, De Involved van de Voorraden op het Consumentengedrag (Amsterdam: Poortpers, 1951), as referenced in, F. Graham Pyatt, ed., Priority Patterns And The Demand For Household Durable Goods (London: Cambridge University Press, 1964), 2.

AN INVESTIGATION INTO THE CAUSAL LINKS
BETWEEN ATTRIBUTION SCHEMA AND DECISION-MAKING

Richard Mizerski, University of Cincinnati
Stephen Green, University of Cincinnati

Abstract

An Interpersonal Replication technique (Bem, 1967) was used to test the hypothesized causal links between perceived attribution schemata and confidence in market information, belief and affect strength, and purchase intention. The results strongly support the validity of a causal link and reveal the differential impact of schema complexity upon consumer information processing.

Marketing has applied the concepts of Attribution Theory in a variety of areas such as children's reactions to television advertising (Robertson and Rossiter, 1974), and advertiser credibility (Settle and Golden, 1974). While the applications may differ, they all probe the general area of how consumers process market information in order to make consumption decisions.

Output of the Attribution Process

The behavioral nature of this process has been suggested by Fishbein and Ajzen (1975), and tested within the marketing environment by Mizerski (1975). These studies suggest that through the process of attributing the causes for events (e.g., product experience, word-of-mouth information, and advertising), beliefs about the stimulus product are formed which then might prompt the development of affect (Mizerski, 1975).

More specifically, Attribution Theory suggests that individuals more readily accept, and are more strongly influenced by, information about an entity if they attribute the report to the entity being described. Applying this to the marketplace, a consumer would perceive that information about a product was accurate and useful the more he believed the content of the information was "caused" by the product being described. For example, suppose the information that a particular brand and model of automobile has good gas economy is perceived to be the result of ("is caused by") the information source actually getting good gas economy with that automobile and telling the receiver about it. If the receiver felt this was the case, he would attribute the information to properties of the stimulus object - a stimulus attribution.

In many cases, however, the consumer may have some doubt about whether the information was "caused" by actual product performance. The receiver may suspect that other causes such as the source's conflict of interest (e.g., the source is selling his own automobile), lack of expertise or specific product knowledge, or some other individual bias (e.g., the source always prefers to speak favorably about the brand) prompted the information. In these latter situations, the receiver makes circumstance (properties unique to the situation or circumstance) and/or person (personal or dispositional properties of the person) attributions, and generally finds the information much less suitable for making inferences about the product.

Measures of Attribution

Research applying these concepts to marketing (see Settle, 1973, and Burnkrant, 1974, for a review) has had a number of limitations. First, causal attributions in marketing research are seldom measured directly. Instead, the strength of the stimulus attribution is often in-

ferred using measures of confidence or credibility. However, research by Newton (1973) and Mizerski (1974) seriously question the validity of this technique.

A second limitation of this research is that it typically elicits the degree of stimulus attribution only, which is not necessarily the lone factor of importance. Consumer perception of what is suitable/useful information is largely determined "... by the configuration of factors that are plausible causes of that information" (Kelly, 1973, p.108). In other words, the type and mix of causes (causal schema) the consumer perceives can also tell us a great deal about how he or she will interpret and process product information.

The Concept of Causal Complexity

In an attempt to overcome these problems, Mizerski (1975, 1977) developed a scale that elicits an individual's causal schema in terms of (1) multiple types of perceived causes, and (2) the relative importance of individual causes in the schema. These values are then transformed into a measure of entropy or an "H" statistic. This transformation provides a measure of the complexity of the subject's causal domain.

As an example, Figure 1 shows five hypothetical individuals' causal allocations among five possible causes for information about a product.

FIGURE 1
HYPOTHETICAL CAUSAL ALLOCATION OF INFORMATION

Perceived Cause	Receiver				
	A	B	C	D	E
(1) Stimulus Attribution	100%			50%	100%
(2) Person Attribution #1		100%		50%	100%
(3) Person Attribution #2					100%
(4) Circumstance Attribution #1			100%		100%
(5) Circumstance Attribution #2					100%

The measure of causal complexity would designate subjects A, B and C the least complex (causally simple), subject E the most complex, and allocations such as D in between.

Theoretical Rationale for the Differential Influence of Causal Schema

The literature (e.g. Mizerski, 1975, 1977) suggests that the five causal schemata shown in Figure One should prompt significantly different cognitive responses in the areas of confidence in the schema, confidence in the information, belief, affect strength and behavioral intention.

Confidence in the causal schema. The literature dealing with the measure of entropy (of which causal complexity is derived) suggests that consumers would be most confident in a causally simple allocation since individuals are less certain/confident in an array with a "broad distribution" than a "sharply peaked" one (Jaynes, 1957). If one considers a consumer's decisions on the number and allocation of cause to be a causal array (see Mizerski, 1975, for a more thorough discussion), an individual should have more confidence in causally simple schemata (receivers A, B, and C) than with the complex schema

represented by E. Also, since we are suggesting schema confidence only at this point, there should be no differences between the three simple schemata on this dimension. H₁: All causally simple schemata will evoke significantly stronger confidence in the schema than a causally complex schema.

Confidence in market information received. A consumer should believe that information about a product was more accurate, and thus have more confidence in that information, the more he assumes that the content of that information was caused by the product being described. Therefore, the less causality attributed to the product evaluated (the stimulus cause), the less confidence he or she will have in the rating information.

H₂: A causally simple stimulus attribution schema will evoke significantly stronger confidence in the product information than causally simple non-stimulus attribution schema (e.g., person or circumstance attributions, receivers B and C)

H₃: A causally simple stimulus attribution schema will evoke significantly stronger confidence in the product information than a causally complex schema (e.g., receiver E in Figure One).

Strength of belief, affect and intention to purchase. In an approach consistent with other researchers, Fishbein (1965, p. 107) has stated that "... a belief about an object may be defined as the probability or improbability that a particular relationship exists between the object of belief and some other object, concept, value, or goal." This definition of a belief appears to be one result of an attribution (Ajzen and Fishbein, 1975; Mizerski, 1975). Kelley (1973, p. 107) notes that attribution is a part of the process "...by which man 'knows' his world, (and) has a sense that his beliefs and judgements are veridical."

It would seem to follow that if an individual made a strong attribution about information concerning product characteristics to a stimulus cause, that would manifest itself in a strong belief that a relationship existed between the product and those characteristics. In Fishbein terminology, the stimulus attribution attaches a higher probability that the information was related to the product. Therefore, the stronger the stimulus attribution, the more extreme the belief. In situations of complex attributions, such as receiver E in Figure One, Kelley (1973, p. 113) suggests that "...the role of a given cause in producing a given effect is discounted if other plausible causes are also present." Therefore, a large number of person or circumstance attributions should discount (in the consumer's mind) the possibility of a stimulus cause.

A substantial amount of research on expectancy - value attitude models suggests that an attitude or affect toward an object is a function of (1) the strength of the individual's beliefs about the object, and (2) the evaluative aspect of those beliefs. While expectancy - value models differ somewhat, each uses some measure of instrumentality or belief strength as a basis for predicting affect. If causally simple subjects form more extreme beliefs, the differential in belief strength should manifest itself in a stronger affect toward the product.

Just as affect has been suggested to be a function of the salient beliefs about a product, a consumer's intention to purchase has been linked to his or her cognitive and affective responses (Lutz, 1977). Although the intention to purchase can be rather distant in both time and number of intervening variables from belief and affect formation, it may also directly respond to differences in causal schemata. Therefore, the following hypotheses will be tested.

H₄: A causally simple stimulus attribution schema will evoke a significantly stronger belief about, affect toward, and intention to purchase the product than

causally simple nonstimulus attribution schemata.

H₅: A causally simple stimulus attribution will evoke significantly stronger belief about, affect toward, and intention to purchase the product than a causally complex schema.

The Question Of Cause And Effect

While the results of previous research (Mizerski, 1975, 1977) appeared to establish a correlational relationship between causal complexity and other processes such as belief formation, there was no evidence that differences in the subjects' attributional schemata caused the differences in confidence, belief strength, etc. In fact, it became apparent that the present paradigm was incapable of testing that question since all the phenomena were cognitive, unobservable behaviors and all were elicited by the same stimulus (i.e., product information). Therefore, even though the present authors strongly believe that the causal schemata preceded and "caused" the beliefs, there was little available evidence to defend that position. How could this question of cause and effect be answered?

A review of the Dissonance Literature, which had undergone a similar problem of investigation that probable causes underlying observed behaviors in experiments, revealed a unique approach developed by Bem (1965). His technique was referred to as an "interpersonal replication paradigm," and it appeared to offer a valid approach for investigating the causal links under study.

Bem's analysis of dissonance, and the present analysis of causal schemata rest upon the single experimental generalization, "that an individual's belief and attitude statements and the beliefs and attitudes that an outside observer would attribute to him are often functionally similar in that both sets of statements are partial 'inferences' from the same evidence..." (Bem, 1967, p.186). Bem gave his subjects information about the treatment a single subject received in a dissonance experiment, and then asked how that subject would have responded. He then argues that when his subjects replicate the dissonance finding, the same process that was operating for his observer was also operating in the dissonance subject. With slight modification, a similar interpersonal replication paradigm was developed for the purpose of testing the preceding hypothesis and to determine whether it seems reasonable to argue that the subjects' causal schemata in the original Mizerski research (1975, 1977) caused the differences in belief and affect.

Methodology

Original Study

In Mizerski's (1975) original experiment, subjects were provided with evaluative rating information on one of three salient attributes for either a fictitious automobile or motion picture. The information was presented in the form of personal ratings that were supposedly made by another individual who was randomly chosen to test and evaluate the product. Following the information treatment, the subjects were asked to allocate the probable cause for the bogus "rater's" opinion (the information treatment) on a causal complexity scale shown in Figure 2. They were then asked to indicate their confidence in the allocation, beliefs about the product, and affective response to the product.

FIGURE 2

How much do you feel that each of the following reasons contributed to the product rater's opinion about the automobile's gas mileage? Note: Any of the following could account for 0% to 100% of the opinion.

- 1. The automobile itself _____%
 - 2. The influence of other people's opinions _____%
 - 3. An effort to please or antagonize the interviewer _____%
 - 4. The personality of the product evaluator (natural tendencies to be critical or complimentary) _____%
 - 5. A general bias for or against automobiles or the brand _____%
- TOTAL _____%

Interpersonal Replication I

Following the basic approach of Bem (1967), an experiment was developed in which subjects (University of Cincinnati undergraduate Business students) would be provided with the same product information as in Mizerski's original study. They were then told to assume that they made one of four possible causal allocations: (1) causally simple to the product - a stimulus attribution, (2) causally simple to a person attribution - "the influence of other people's opinions," (3) causally simple to a circumstance attribution - "the personality of the product evaluator . . .," or (4) a causally complex attribution which provided approximately equal allocation to each possible cause (refer to Figure 2). Ninety four percent was allocated to each simple attribution, with one or two percent allocated to the remaining causes. In the complex treatment, between 18 and 22 percent were allocated to each of five causes.¹

After being provided with the rater's information that the automobile tested rated "superior" on gas mileage and the treatment causal allocation, the subjects were asked to respond to several questions. The first question asked "How much confidence would you have in your assignment of percentages for the product rater's opinion about the automobile's gas mileage" (confidence in the schema). The second question asked "How confident do you think you would be that the product rater's opinion of the automobile's gas mileage was accurate; i.e., that the automobile actually had superior gas mileage?" This measured the subjects' confidence in the rater's opinion. The subjects' responses were gauged on a nine point scale that ranged from "No confidence" (#1) to "Complete confidence" (#9).

They were then asked a question about the strength of their belief that the automobile had superior gas mileage (unlikely--#1, to likely --#9), and the strength of their affect toward the stimulus product (extremely low appeal --#1, to extremely high appeal--#9). Finally, they were asked, " . . . If you felt that an automobile's gas mileage was very important, what do you think is the probability that you would actually consider purchasing this automo-

¹Originally, the simple causal allocation treatments had the total 100% allocated to the relevant simple cause, with an equal 20% allocated to each of the five potential causes in the causally complex treatment. However, pretests showed that the subjects inferred that the causal allocator " . . .did not take enough time" with the task using those percentage allocations.

bile if you were in the market when it is introduced?" Their response was scored on a scale that ranged from "very low probability" (0%) to "very high probability" (100%). All hypotheses were tested with a one-way analysis of variance, with each replication tested separately.

Interpersonal Replication II

After the initial interpersonal replication experiment was completed, it was felt that the experiment should be expanded so that results from treating all the potential simple causal schemata could be examined. Therefore, a second experiment treated the following allocations: (1) causally simple to the circumstance - "An effort to please or antagonize the interviewer;" (2) causally simple to the person - "A general bias for or against automobiles;" (3) the same causally simple to the product; and (4) causally complex schemata provided the subjects in the first interpersonal replication study.

The first interpersonal replication was run on two undergraduate marketing classes; the second experiment used two undergraduate management classes. Both samples closely approximated the backgrounds of the individuals in Mizerski's original study. The treatments were randomly assigned to the subjects, with each subject exposed to one of four treatments in each experiment. Because of the similarity in methodology, the results of the two studies will be discussed together in the text.

Results

Confidence in the Schema

In an effort to simplify presentation of the data, the results of both experiments are provided in Table I. Standard deviations of the mean scores are shown in parentheses, with cell sizes reported to the left.

The first hypothesis proposed that there would be no significant differences between the simple causal schemata, that they would all prompt stronger confidence in the schema than the complex schema. The mean scores from both studies (the first column of Table I) show only partial support for the hypothesis. The simple attribution to the product was the only simple schema that prompted significantly stronger confidence than the causally complex allocation ($t=2.34$, $df=46$, $p<.01$; and $t=3.43$, $df=58$, $p<.001$ for study I and II, respectively).

It was found that the simple attribution to the product also prompted significantly stronger confidence in the schema than the other simple attribution ($t=3.20$, $df=46$, $p<.001$, and $t=5.64$, $df=58$, $p<.001$). This latter finding was not expected since the literature strongly suggests that the more simple the attribution, the more confidence the attributor should have in the schema (i.e., a simple attribution is a simple attribution for this dimension).

TABLE 1
MEAN SCORES FOR DEPENDENT MEASURES

Treatment/Causal Attribution	Confidence in the Schema *	Confidence in the Opinion *	Belief Strength*	Affect*	Intention**
Experiment I					
11 Simple to the product	6.46 (1.76)	6.62 (1.45)	6.23 (2.06)	7.11 (2.04)	69.23 (26.91)
11 Affect to please interviewer	4.69 (2.09)	3.45 (1.86)	3.64 (1.43)	5.36 (2.61)	57.27 (19.62)
11 General bias for or against	4.31 (2.66)	3.85 (2.15)	4.31 (2.01)	6.09 (2.19)	56.39 (24.36)
13 Complex attribution	5.54 (1.66)	4.62 (1.94)	4.92 (1.65)	6.36 (1.19)	62.23 (20.72)
Experiment II					
13 Simple to the product	6.27 (2.21)	6.53 (1.88)	6.29 (2.21)	7.40 (1.24)	61.33 (24.16)
13 Influence of other's opinions	3.27 (2.28)	3.13 (1.81)	3.40 (1.50)	5.67 (2.06)	46.00 (27.62)
16 Personality of rater	2.56 (1.99)	2.81 (1.27)	3.06 (1.23)	5.44 (2.22)	45.00 (28.75)
16 Complex attribution	3.94 (1.76)	3.18 (1.32)	3.63 (1.15)	4.61 (1.90)	45.00 (21.99)

* Scale ranges from 1 (low value) to 9 (high value).

** Scale ranges from 0% (very low probability) to 100% (very high probability).

Confidence in the Accuracy of the Rater's Opinion

The second hypothesis suggested that the subjects would have more confidence in the accuracy of the rater's opinion when using schemata reflecting simple allocations to non-product causes. The mean scores in both studies are in the appropriate direction, (See Table 1) and the results of planned comparisons ($t=4.60$, $df=46$, $p<.001$; and $t=7.13$, $df=58$, $p<.001$) support the hypothesis.

The third hypothesis, that subjects who received the simple causal allocation to the product, would have more confidence in the accuracy of the rater's opinion than those who received a causally complex schema, was also supported ($t=2.73$, $df=46$, $p<.005$; and $t=5.86$, $df=58$, $p<.001$).

Strength of Belief, Affect, and Intention to Purchase

The fourth hypothesis proposed that subjects receiving a causally simple stimulus attribution schema will produce stronger belief, affect, and intention to purchase mean scores than subjects receiving causally simple person or circumstance schemata. Hypothesis Five went on to suggest that the same differential strength of the simple stimulus attribution schema would also be reflected over those subjects who received a causally complex schema. The easiest way to evaluate the results is to discuss one element at a time, across the two experiments, since significant differences were observed.

Both hypotheses were supported in the two experiments for the measure of belief strength. Subjects receiving the simple stimulus attribution schema formed significantly stronger beliefs than subjects under the simple circumstance and simple person treatments ($t=3.57$, $df=46$, $p<.001$; and $t=5.18$, $df=58$, $p<.001$), and stronger than subjects receiving the causally complex schema ($t=1.82$, $df=46$, $p<.037$; and $t=3.93$, $df=58$, $p<.001$).

As noted earlier, the formation of beliefs has been discussed as a direct response of the attribution process.

On the other hand, the development of affect and intention to purchase is a partial function of the beliefs about the stimulus object. One would expect somewhat weaker relationships with the attribution schema used.

As hypothesized, the subjects receiving the simple stimulus attribution schema formed significantly stronger affect than subjects with simple non-stimulus schemata ($t=2.31$, $df=46$, $p<.013$; and $t=3.11$, $df=58$, $p<.002$). However, the prediction that a simple stimulus attribution schema would also prompt stronger affect than a causally complex schema was not fully supported in both experiments. While the mean scores in the two experiments were in the appropriate directions (see Table 1), they failed to reach significance ($t=1.15$, $df=46$, $p<.128$) in experiment #1, while there was strong support for the hypothesis ($t=3.81$, $df=58$, $p<.001$) in the second experiment.

Similar results were found for the subjects' intention scores. Although the mean scores comparing the simple stimulus schema to other causally simple non-stimulus schemata (hypothesis 4), and to causally complex schema (hypothesis 5) were in the appropriate direction, hypothesis four was supported marginally ($t=1.43$, $df=46$, $p>.079$) and hypothesis five was not supported ($t=0.77$, $df=46$, $p>.44$) in the first experiment. Both hypotheses, however, were supported in the second experiment, ($t=1.81$, $df=58$, $p<.038$, and $t=1.74$, $df=58$, $p<.044$).

Summary And Conclusions

The results strongly support the hypothesis that differences in the complexity of causal schemata can cause differences in other cognitive responses. Specifically, those individuals using a causal schema that reflected simple allocations to the product (stimulus attribution) responded with significantly higher levels of confidence in the accuracy of the information they received, stronger beliefs that the product possessed the characteristics assigned it by the information source, and a stronger affective response to the product than individuals using a complex causal schema or schemata reflecting a simple non-stimulus attribution. In addition, those subjects using a simple-product schema demonstrated a stronger intention to purchase the product than subjects using a simple-circumstance or simple-person attribution schema.

The comparison of simple-product and complex schemata on intention to purchase, however, yielded equivocal results. Although it is not clear why the subjects using the simple-product schema formed stronger intentions to purchase than subjects using a complex-schema in the second study only, one could speculate that several factors might have been operating. First, as previously noted, intention to purchase is only partially a function of beliefs and therefore not as directly related to schemata. Thus, it may be that this phenomenon is more tenuous and difficult to capture experimentally. Also, one should note that all the findings are stronger with the second study than in the first. This discrepancy may be due to the fact that Study One was conducted in marketing classes where the students were more familiar with marketing concepts and research, and therefore may have been "test wise" and wary in their responses.

The failure of hypothesis one, that all simple schemata will evoke stronger confidence in the schema than the complex schema, may have been due to the nature of that question more than any "real" psychological process. In pretesting the instruments, it was found that the subjects had a great deal of difficulty understanding the intent of the question concerning confidence in the schema. After many iterations on that measure, it was felt that it had been made clear. Debriefing, however,

revealed that subjects had confused the confidence in the schema measure with the confidence in the accuracy of the rater's opinion measure (these two responses correlated .54, $p < .001$, when the treatment means were removed from the scores). Therefore, any interpretation of the confidence in the schema data appears problematic.

Nonetheless, it is apparent that differences in causal schemata complexity can cause differences in other important cognitive processes. To the extent that the process revealed in our "observers" is the same process that occurs in individuals who generate their own causal schema (as Bem, 1967, and these authors would argue), there are clear implications for marketing strategy. If one can devise advertising formats that elicit a simple-product causal schema, as opposed to simple nonproduct or complex allocations, the consumer should be more likely to be confident that the product information is accurate, as well as form stronger beliefs about the product's characteristics. Similarly, the consumer should have a stronger affective response toward that product and might (but not necessarily) develop a stronger intention to purchase the product.

Research is presently underway in an attempt to identify the varying capacities of advertising formats to elicit a simple-product schema in consumers. Other research might focus on the effects of different media presentations upon the complexity of evoked schemata or investigate the possibility of causal complexity existing as an individual difference construct.

REFERENCES

- Bem, D.J. "An Experimental Analysis of Self-Perception," Journal of Experimental Social Psychology, 1 (1965), 199-218.
- Bem, D.J. "Self-Perception: An Alternative Interpretation of Cognitive Dissonance Phenomena," Psychological Review, 74 (1967), 183-200.
- Burnkrant, R.E. "Attribution Theory in Marketing Research: Problems and Perspectives," in M.J. Schlinger (Ed.), Advances in Consumer Research, Vol. 2, Proceedings of the Association for Consumer Research, 1974.
- Fishbein, Martin. "A Consideration of Beliefs, Attitudes, and Their Relationship," in Ivan D. Steiner and Martin Fishbein, eds., Current Studies in Social Psychology. New York: Holt, Rinehart, and Winston, Inc., 1965.
- Fishbein, Martin, and Icek Ajzen. Belief, Attitude, Intention and Behavior; and Introduction to Theory and Research. Reading, Mass.: Addison-Wesley, 1975.
- Jaynes, E.T. "Information Theory and Statistical Mechanics," Psychology Review, 106 (1957), 620-631.
- Kelley, Harold H. "The Processes of Causal Attribution," American Psychologist, 28 (1973), 107-128.
- Lutz, Richard J. "An Experimental Investigation of Causal Relations Among Cognitions, Affect, and Behavioral Intention," Journal of Consumer Research, 3 (1977), 197-208.
- Mizerski, Richard. "A Test of the Relationship Between Trait and Causal Attribution," in J. Schlinger, ed., Advances in Consumer Behavior, Vol. 2, Proceedings of the Association for Consumer Research, 1974.
- Mizerski, Richard. "An Investigation Into the Differential Effects of Causally Simple and Complex Attributions," in B. Anderson, ed., Advances in Consumer Behavior, Vol. 3, Proceedings of the Association for Consumer Research, 1975.
- Mizerski, Richard. "Causal Complexity: A New Measure of Consumer Causal Attribution," Working Paper, University of Cincinnati, 1977.
- Newtson, Darren. "Attribution and the Unit of Perception of Ongoing Behavior," Journal of Personality and Social Psychology, 28 (1973), 28-38.
- Robertson, Thomas and John Rossiter. "Children and Commercial Persuasion: An Attribution Theory Analysis," Journal of Consumer Research, 1 (1974), 13-20.
- Settle, Robert and Linda Golden. "Attribution Theory and Advertiser Credibility," Journal of Marketing Research, 11 (1974), 181-185.

SOME ISSUES IN DESIGNING CONSUMER INFORMATION STUDIES IN PUBLIC POLICY

R. Bruce Hutton, University of Denver
Dennis L. McNeill, Arizona State University
William L. Wilkie, University of Florida

Abstract

In order for the consumer researcher to make significant inputs for public policy decisions regarding information provision, he must be aware of the nature of the policy environment and the kind of research that will best fit its needs. This paper addresses several issues of importance in doing policy related research and describes two specific studies in the area of energy information undertaken at the University of Florida's Center for Consumer Research.

Introduction

The content and availability of product information for use in evaluation and decision making by consumers is a major topic of interest and concern among various groups -- marketers, policymakers, consumer educators, and consumer researchers. At issue is the availability and subsequent use of information by which consumers can make more "informed" choices (e.g., objective information representing specific physical features). Programs to provide such information have been created in a broad range of government agencies, including the Environmental Protection Agency, Department of Transportation, Food and Drug Administration, and Federal Trade Commission. Current initiative is reflected in the 1975 Energy Policy and Conservation Act (EPCA) whose basic mandate is to "conserve energy supplies through energy conservation programs . . ." (EPCA, 1975). One of the major priorities of this program is the establishment of labeling rules for major durables, specifically in the area of energy consumption. For the consumer researcher to provide relevant input to such decisions, his research paradigm should reflect the special problems in policy oriented research.

Trade Regulation Rule Influence

Day (1976) notes that we are certain to see an increase in the disclosures of efficiency performance data in the future. The primary vehicle for operationalizing future information provisions will likely be through the Trade Regulation Rule (TRR) power of the Federal Trade Commission (FTC). In effect, TRR's allow a public agency to require particular marketing practices to be performed by all marketers in a given product or service category (Wilkie and Gardner, 1974; Wilkie 1975c). This approach by the FTC will require policy decision inputs at a number of different stages. Wilkie (1975) has delineated six stage process of TRR development:

- (1) Selection of product classes.
- (2) Identification of relevant product characteristics.
- (3) Development of standards and test methods.
- (4) Determination of reporting format.
- (5) Provisions for dissemination.
- (6) Assessment of effectiveness.

The different stages represent different levels of importance for the consumer researcher. Stage 1 is crucial because priority concerns and program objectives are formulated here. Stages 2 and 3 will be primarily undertaken by technical experts. Stages 4, 5 and 6 provide the major opportunities for consumer research.

There are both major and complex questions about the

need for TRR's, especially in a cost/benefit sense. Expected benefits from TRR's can be classified into two major types: (1) administrative/legal benefits, and (2) economic/social benefits (Wilkie, Hutton, and McNeill, 1975). TRR's are expected to make the regulatory process more efficient by removing the need for case by case adjudication and to provide marketers with better guidelines for what is expected of them. In the second case, TRR programs should bring the system more towards the idea of the "fully informed consumer" and to give him more objective information for product/brand decisions. At the time of this writing, TRR proposals aimed at increasing consumer information were underway in a number of areas, including: nutritional information in food advertising, OTC drug advertising, used automobiles repair information, standardized product warranties, comparative price disclosures for life insurance policies, detergent cleaning ratings, and energy consumption labels for household appliances. Each of these areas involves millions of consumer purchases and potential for considerable improvements. Consumer research is obviously relevant for program decisions.

Purpose of this Paper

The design and conduct of research for public policy needs has been a focus of attention at the University of Florida Center for Consumer Research. A series of experiments have been conducted which deal specifically with the area of consumer product information. In designing these studies, the authors have encountered several important issues of likely interest to both the academic researcher and policymaker. The purpose of this paper is to examine these issues, show how they have influenced two particular studies dealing with consumer impact of "new" product information, and to provide some general suggestions for others' consideration of this research area.

Preview of Issues

In any research project, decisions will have to be made regarding the focus of the research, design issues, data collection, and analysis. In policy related research, while some issues will be unique to the public policy field, the majority of decisions to be made will be of the same type as any research project. However, the final approach the researcher takes will be different in many cases because he or she is addressing the research to policy needs. The subsequent policy-related approach carries with it a series of issues and problems, the combination of which is not faced in the normal research program. Major areas of concern for such policy related research are presented in Table 1.

The Role of Policy Objectives

In any research setting, it is crucial for the researcher to have the focus of the study clearly in mind. In the case of research applied to policy needs, an understanding of policy objectives is a necessary condition to defining the problem and establishing the reasons for investigation.

Wilkie, et.al., (1975) state that vague or unspecified objectives by the policymaker can easily allow differing research thrusts, designs, and measurements supposedly aimed at the same policy decisions. The risk is that

TABLE 1

Some Public Policy Research Issues	
Issue	Areas of Concern
1. Role of Policy Objectives	-- Problem definition -- Basic v. applied research -- Stimulus -- Unit of analysis -- Criteria
2. Research Design	-- Design format (task orientation) -- Validity -- Demand characteristics
3. Data Collection	-- Focus groups -- Measurement and scales -- Sampling
4. Data Analysis	-- Levels of significance -- Statistics

some of these will be inappropriate for the policy or program mandate. Consider, for example, the fundamental problem of determining dependent variables for studying the effect of information on consumers. A natural tendency would be to choose brand choice or at least attitude and intentions as appropriate variables. But these may not be appropriate criteria. For example, if the goal of the recent energy program is to reduce demand, this would not be reflected in brand shifts as much as model shifts. This point is also made by Bettman (1975) who states that the choice of criterion variables is a function of intent of the information provision. If the intent is that the information be used in a specific manner, attitudes and behavior become more important relative to recall, knowledge, and awareness. If the intent is that consumers should be aided in perceiving and processing the appropriate information but no commitment concerning their use of it, recall, knowledge, and awareness variables become crucial.

Once the researcher understands the basic problem, he must define for himself the rationale for investigating the area (e.g., timely, theoretical significance, practical significance, instrument development, clarification of existing perspectives, stream of research). Based on the reason(s), the decision to be made is whether the focus of the research will be problem or theory oriented. Wilkie (1975) addresses this issue in his discussion of basic v. applied research. The criteria for choosing may be established as:

<u>Applied</u>	<u>Basic</u>
<ul style="list-style-type: none"> • primarily responsive to and reflective of the policy setting • results offer useful perspectives to the policymaker and marketer • enhance knowledge of the way consumers perform 	<ul style="list-style-type: none"> • derived from theory • tight testing of hypotheses • lack of direct short run benefits to the policymaker

Once the researcher has decided these fundamental questions, more specific issues arise concerning questions of appropriate stimuli, units of analysis, and criteria.

Should the researcher choose to focus on the programs, the stimulus chosen for study should meet a set of criteria based on policy needs. The product class chosen would be a function of: (1) the state of knowledge of the consuming public and subsequent sample, (2) characteristics of product use, and (3) the presence of within product class variance. That is, the product class should show variability along the dimension to which

the program is directed.

The unit of analysis in addressing policy questions is usually the individual as opposed to the product. Some experiments done in the information provision area focus on shifts in choice (e.g., Houston, 1972; Isakson and Mauizi, 1973; Russo, et.al., 1975); therefore, the unit of analysis is necessarily brand, which does not measure information use directly. For most information programs however, policy-oriented research will focus on the individual consumer level at early stages.

In the past, criteria for evaluating the effects of information have tended to be market share or brand switching behavior. The appropriateness of utilizing these measures alone in evaluating information provision can be questioned, especially when one recognizes that the consumer is faced with a new environment and brief exposure to the information. In such a limited exposure, the use of brand choice places a great deal of responsibility on this level of impact. As with other marketing research, it is more advisable to measure the effects of information at different levels of response using multiple dependent measures (Day, 1976; Heeler and Ray, 1972). Multiple dependent measures will most likely require multiple tasks, with emphasis on realistic content and order.

Research Design

The choice of the appropriate experimental design is, of course, dependent on a number of factors. However, in policy related research, stress is placed on external validity.

The key, then, is to make the experiment more valid externally without losing internal validity. Only two experimental designs meet the criteria for internal validity and control at least one source of external invalidity -- Solomon four-group design and posttest-only control group design (Campbell and Stanley, 1963). Because of the cost involved in running the Solomon-four, the posttest only is usually preferred. Both of these designs control for the possible interaction of testing and X (i.e., the experimental treatment). However, two other interactions are also a problem. It is possible that any effects demonstrated hold only for that unique population from which the groups were selected (interaction of selection and X). In order to reduce this problem the researcher should increase the number and types of groups used. The other interaction involves the effects of recent events on subject responses (interaction of history and X). For example, an experiment on energy information done immediately following the presentation of President Carter's energy program might produce a responsiveness to X not to be found on other occasions. Consequently, it is desirable for policy research to lead, not lag, the phenomena under study.

The other major concern for the researcher in this area involves demand characteristics. For the policy researcher two factors are especially relevant. First, the subject is placed in a changed information environment. Consequently, exposure to new concepts may cue him in terms of the focus of the study. Second, the experiment set up to have the subject engage in multiple tasks is open to demand characteristics if one task cues subject response on other tasks in ways not intended by the researcher. Consequently, the flow of tasks and measures is a primary consideration for the researcher. For an excellent summary of demand characteristics in lab experiments and how to deal with them, see Sawyer (1975).

Data Collection

Focus group interviews can play an important part in the development of the policy research format, with primary

purposes being to: (1) Explore the current state of consumer knowledge and attitudes toward the stimulus; (2) Obtain a better understanding of the consumer language used in thinking about the stimulus.

The results of the focus group interviews will provide important information for the development of scales for subject response. Scale construction should be a function of the population under a study and not the desired analysis procedures. Obviously scales and tasks should be set up so they can be understood by all subjects. Unfortunately, this may not always be possible, especially when dealing with some segments of the population. In order to minimize this problem, scales should reflect as much as possible the language of the consuming population and real world choices.

Other important questions in the research design involve sampling procedures. In order to provide policy guidance, it is obviously desirable to use subjects which reflect a realistic population of consumers for the program at issue. Lab experiments will normally not be feasible with probability samples, but neither are they normally very useful with college sophomores. One guide might be to use real consumers, attempting to include a broad range of demographics within the sample.

One economical (time-wise) approach is to solicit "intact groups" such as social, religious, and educational organizations. (This will, however, often require within-session assignments to different conditions, which will impact on the design and operationalization of the study itself.) If this approach is taken, the researcher should vary the nature and demographic character of the groups chosen, use as many groups as possible, and have no single group dominate the sample.

Real consumers, while providing more relevant data, do present some formidable design problems when compared to college students. Special care must be taken to prevent difficulties the subjects might encounter with certain abstract concepts and with a potential reactive response to being placed in a "testing" mode. These problems require careful development of instructions and tasks and attention by those involved in the conduct of the experiment. Some guidelines we have developed based on recent experiences include:

- Provide complete assurance regarding the anonymity of subject responses. This will help reduce the feeling of being tested and graded.
- Researcher should be ready to accept the fact that some segments of the population just may not be able to do all tasks. This may reflect the study's demands rather than the program being studied; it may be necessary to do a separate study for these segments.
- Non-responses are a fact of life. Some subjects will refuse to make a "best estimate" or "guess" because they feel that would be dishonest or a form of cheating.
- It may be beneficial to provide the subjects with familiar surroundings (i.e., go to their church, club, etc.). Since the researcher is now out of his familiar setting, he should be ready to meet a variety of unforeseen occurrences. Examples these authors encountered include a subject's breaking into tears, a blind subject, mothers missing directions because of a 5 year old's nature call, and confrontation with an inebriated by-stander. Dealing with these kinds of factors certainly approaches art more than science.

Data Analysis

The relationship between the issues being discussed is not one of independence. Naturally, many of the factors discussed previously will affect the data analysis stage (e.g., non-responses, scale construction, etc.). However, there are several specific points that should be advanced:

- The type of people and responses they might normally make will influence the analysis procedure. For example, in the construction of appropriate scales, the result may, in the interest of realistic alternatives, be a nonmetric, as opposed to metric scale.
- There is a need to determine in advance the accepted level of significance. Typically is set at the "sacred" .05 or .01 level. In fact, a number of criteria should be used to determine the appropriate level (Labovitz, 1970). In dealing with issues and samples of the paper above, less stringent significance levels may be justified. In any case, it is helpful to report the actual level attained and overtly provide the researcher's judgment as to the meaning of the obtained result.
- In the case of multiple tasks and measures, analysis should be done in light of other responses. For example, a complexity of information measure was taken on a standard seven point bipolar scale. The results showed no difference across groups with all groups judging it "simple." A clear case of easy to understand information? Subsequent tasks involving the use of the information showed the groups acted differently, and suggesting that many subjects, in fact, may not have recognized the complexity of the information and only dealt with it in a superficial manner. Additional items and analyses can be included to address these possibilities.

Issues Specific to Energy Labeling

In light of recent developments (e.g., CIA report on energy resources, President Carter's energy program, development of a Department of Energy), issues of the effects of energy labeling of appliances on consumer response have become particularly relevant.

Energy Programs

The Voluntary Labeling Program resulted from a Presidential directive in April 1973. This program was established within the Department of Commerce to deal exclusively with major energy consuming household appliances. Its general purposes were:

- (1) Encourage manufacturers (including private brand labelers) to voluntarily provide consumers with information concerning the energy efficiency or energy consumption of major durables.
- (2) Encourage consumers to utilize the information when evaluating products by providing, at the point of sale, energy information presented in a uniform manner and readily understandable in order to facilitate product comparisons.

Early success included the fact that 24 manufacturers representing an estimated 95% of room air conditioner sales in the U.S. were participating in the program as of 1975 and generally favorable consumer attitudes were

exhibited toward such labels. In 1976 this voluntary program was abolished by the Energy Policy Conservation Act, and was replaced by a mandatory labeling program under the primary jurisdictions of the FEA and FTC.

The Energy Policy and Conservation Act (EPCA) passed in December 1975, reflects a broad range of programs. Included as important thrusts are efforts on product testing, labeling, energy standards, and consumer education. The Act requires that labels disclose, for each model, (1) the estimated annual operating cost and (2) the range of alternative's estimated annual operating costs. In addition the FEA has the power to prescribe energy efficiency standards to any type or class of covered products. However, it is stipulated that resulting energy savings must outweigh any increase in price or maintenance expense, lessening of product performance, or negative effects on competition in order to be implemented. Finally, the FEA will also carry out a program to educate and encourage consumers to use such information in conserving energy.

Disclosure of Energy Use Data: Study A

The purpose of this research was to address and examine certain consumer information processing issues within the Energy Act's formal programs of standardized energy consumption disclosure. The issues surrounding the energy labels are important, in that the form and content of the labels may well effect the impact of the program on consumers. This research focused on two central questions:

- (1) Will the presence of energy consumption data on labels attached to major home appliances be likely to assist the consumer in purchase related activities?
- (2) Is the reporting format required under the Energy Act the best way to disclose energy consumption?

The first question is important from a program evaluation perspective, and, in addition, will provide a frame of reference for the investigation of the format question. The impact of the disclosure, given the early stages of program development and the timing of this research, should provide insights into the potential impact of the program and whether it appears wise to proceed.

The second area of interest is the format of the disclosure, an area which appears to have been infrequently researched. Particular issues of interest regarding the energy labeling format were:

- Unit of Measurement -- The required disclosure is in dollars. The unit of measurement question is important in light of program goals. For this study the units of measurement under investigation were the dollars required by the EPCA and the kilowatt per hour (KWH) alternative.
- Comparative Disclosure -- The EPCA requires a new dimension of product performance when it requires comparative energy use for similar products. Effect of this additional information was investigated.
- Time Period for energy use computation -- The required time period is in annual terms. The issue is one of magnitude, for the time period used can inflate the model consumption figures. For this study a comparison of monthly and yearly time periods was conducted.

• Degree or Specificity of disclosure -- A crucial aspect of the disclosure of energy use is the ability of the consumer to identify those attributes of the products which account for energy use. The impacts of explicit -- by feature -- disclosure of energy use were also assessed in the study.

Energy Use Data/Policy Research Issues

The combinations of each of these issues required too many experimental cells for the limited resources of the study. Rather than eliminate controllable decision variables, it was decided that certain combinations should be dropped, these determined in light of that combination's realistic chances of actually being used by the program. Readers will note that this decision incurs costs in terms of our subsequent ability to disentangle main effects of the format variables mentioned above. On the other hand, it does allow us some assessment of more variables than would otherwise have been possible. Six conditions were run in the study, ranging from control, with no energy use information, to annual dollars with a range, broken down by feature. The design was completely randomized, after-only, with a control condition and five disclosure conditions.

The product class chosen, refrigerator-freezers, is an interesting (and relatively straightforward) one from the program perspective. The product occupies a prominent place in the home and accounts for a significant proportion of average household energy use. The use of the product is relatively constant across households, reducing the variation in actual consumption as a function of household and geography. In addition, the variability in energy utilization is predominantly associated with features, rather than only brand, choice.

In determining the tasks and measures, the program's objectives played an important role. The goals of this particular program are overtly to reduce the demand for energy. One research focus thus should be on consumers' behaviors re: energy. However, as noted earlier, features are more significant than brands here, such that research tasks had to be altered from the normal stress on brand preference or choice. To assist in more detailed decisions, focus groups were conducted, providing inputs for the language and attributes included in stimulus development.

Table 2 charts the conduct of the experiment. The purpose of the experimental tasks was to allow for a complete profile of energy use impact through multiple dependent measures, while the subjects encountered the information in realistic situations. This set of tasks and measures took slightly over one hour; a fuller description is available in McNeill (1977).

Essentially, the tasks reflected three stages of information exposure and use. The initial exposure required the subject to acquire the information (ala Jacoby's "information board") by pulling tabs attached to four individual model displays. (Models had been designed to allow research inferences as to how well the information was being communicated and used.) In this way, the subjects learned about the models and the researcher could follow "where the subject's eyes go" when evaluating the model. The subjects were then asked to evaluate the models based on unaided recall of the information they had acquired. The second level of information exposure was a second visit to the same models, now allowing external memory while pursuing the evaluation tasks. The final level of information effect was a unique operationalization of the choice dimension. This requested the subject to "build" their own personal model of refrigerator-freezer feature by feature, with a number of subsidiary measures taken afterward. This series of experiment tasks provided responses to the in-

formation across a wide range of impacts, allowing a reasonably comprehensive analysis of effects. Results and conclusions are detailed in McNeill (1977).

Results

Given the comprehensive design of this study with its multiple measures and tasks, a detailed explanation of findings is beyond the scope of this paper. Overall, the results of this study should prove useful to both policymakers and consumer researchers. The presence of energy consumption data had a significant impact on consumer's information processing responses. After the initial shopping task (involving information acquisition) the product preferences, overall impressions, and energy use judgments of the experimental subjects were in the hypothesized direction. This was also found in evaluative judgments in the presence of the information on the second shopping trip task. Essentially, the nature of the results shows that the higher the energy use the lower the model evaluative judgments. This is supportive of the energy use impact potential in light of both the intent and content of the disclosure.

The second question under study concerned the format of the information and the relationship of format to consumer information processing response. The format variable which produced differential response was the comparative disclosure. Given that the goal of this disclosure was to provide more effective relative judgments of model performance, a set of specifically designed scales were developed. The comparative judgments reflected the comparative disclosures (in the case of relatively high energy use) on the likelihood of finding a better value, the worth of additional shopping, and purchase recommendations. While the other format variables did not produce significantly different responses, there were no format variables which appeared to hinder the consumer response in terms of the goals of the disclosure. The format question, however, is an important one and must be investigated from other perspectives in order to make completely profile this decision for the policymaker.

Life Cycle Cost: Study B

While present policy consideration is focused on the format questions discussed above, it may be that a more comprehensive information scheme is much more meaningful. This index, termed "life cycle cost," provides a framework for the summary presentation of three product dimensions -- price, energy, and service costs (M.I.T., 1974a). Life cycle cost (LCC) is defined as the discounted sum of all dollars paid for an average product during its useful life (i.e., purchase price + energy cost + service cost). In essence, LCC provides the consumer with an organized and consistent way of processing all product cost information by incorporating the three cost dimensions within a consistent time frame (i.e., average product life) and common units of measurement (i.e., dollars).

Our study was aimed at exploring the impact of LCC, as a new information form, on consumers. Specifically, the study compared the existing consumer information environment -- with its emphasis on purchase price -- to a new environment providing LCC data. In addition, the study stressed multiple tasks and measures aimed at more comprehensive evaluation of information effects, attempting to account for the fact that such a program would not have been encountered before.

M.I.T. Report

Background for, and development of, the LCC concept is described in the monograph "Productivity of Servicing Consumer Durable Products," authored at M.I.T. under an

TABLE 2

Experimental Design Flow Charts for Two Studies

Disclosures of Energy Use Data

Focus Group Interviews

Introduction

- Purpose of experiment
- Release form
- Present concern for refrigerator-freezers

Initial Model Evaluation

- Model information acquisition
- Model preferences
- Overall model impressions
- Feature recall

Model Evaluation (non-recall)

- Model evaluation by feature
- Comparative advice

Building Task

- Feature decisions
- Feature/energy use judgments

Demographics

Choice of Energy Use Disclosures

Life Cycle Cost

Focus Group Interviews

Subject Release and Cover Story (assurance of anonymity)

Generic Background Data on Refrigerator-Freezers

Model Evaluation (information by models matrix)

- Absolute satisfaction scores based on needs and budgets
- Relative comparisons between models based on needs and budgets

Subjects "Build" Their Own Refrigerator-Freezers

- Written protocols concerning how and why they choose the features they did
- Subjective state questions concerning the certainty, helpfulness, and complexity of choices and information

Feature Evaluation

Price and Energy Cost Perceptions (absolute amounts and comparative)

Estimate of Feature Costs (recall and recognition)

Calculation Tests

Demographics

Energy Information Preferences

NSF-RANN grant. The report asserts that consumers fail to recognize energy costs in appliances, do not seek out energy and service related information during product evaluation, and instead opt for increasing use convenience at the expense of increasing energy consumption. Since higher prices may be necessary for the introduction of energy conserving design changes, a shift in consumer perceptions of cost from sole reliance on initial purchase price to a long-run total (LCC) cost appears necessary for the market to operate so as to

reduce energy consumption. The nature of presumed consumer impacts from LCC information is presented in Wilkie and Hutton (1977).

LCC Policy Research Issues

The combination of a new conceptual framework for cost and a policy orientation provided this study with a number of challenging issues. First, the introduction of a new concept such as LCC provides the consumer with a different environment in which to operate. Multiple tasks were needed to reflect different levels of consumer response expected within the proposed new environment. Index figures for LCC had to be realistic, which required technical advice and a series of product cost calculations on refrigerator-freezers, which were chosen for this study for much the same reasons as noted earlier. Also, major concern was again with factors in the product that relate to cost (i.e., features). In the case of refrigerator-freezers this would include size, frost free, automatic ice maker, power saving switches, etc. Stimuli are identified as models (combinations of features) since they reflect the different cost combinations more so than brand by brand comparisons of the same model types.

The research design was a posttest-only control group design. Besides the LCC condition, an energy/year condition was included to reflect the EPCA mandated program. Both experimental groups represent a changed environment from today. Intact groups were used, so random assignment to conditions was done within groups instead of among the total subject population.

A flow chart of the experiment is given on the bottom of Table 2. Design of the tasks and measures involved some difficult trade-offs among policy concerns, research control, analytical needs, and subject's capabilities and motivations. Some examples: (1) how many models to present? (2) whether, and how, to introduce budget constraints, (3) whether scales should be balanced versus skewed when dealing with arguably positive or negative concepts, (4) how to assess basic capabilities (e.g., computation skills) without overtly testing, (5) how to represent negative costs (savings), (6) how to introduce this new information together with some education on it, while minimizing demand characteristics, and (7) determination of appropriate criteria by which to reach summary conclusions, given the large number of measures. Details of these decisions are available in Hutton (1977).

Results

Once again, given the extensive design of the study with its multiple measures and tasks, a detailed explanation of findings is beyond the scope of this paper. For a complete analysis of results see Hutton (1977). In general, however, the strongest results were seen in tasks reflecting levels of consumer response in a more cognitive as opposed to behavioral sense. The study provides support for the provision of more comprehensive objective information. By letting control group responses represent what consumers know going into a buying situation and LCC responses what is learned, the differences indicate that consumers do learn from the presentation of such information. In fact, control group responses show a significant underestimation of the magnitude of operating costs. There is also evidence that what is learned by the consumer is more than a specific cost figure but rather a different conceptual view of product cost from the one dimensional association with price toward a multidimensional long run concept.

From a policy standpoint, it is important to know the impact of a proposed provision along various dimensions of utilization. Can they utilize the information correctly or does it result in confusion, misinterpreta-

tion, or ignoring of the information package? Findings along these lines are equivocal. Consumers do utilize some aspects of the information, even lower educated ones, but not without certain negative consequences. In addition, there is some indication that consumers will evaluate models and features which conserve energy more favorably and energy consuming features less favorably with the availability of energy cost data. In terms of purchase behavior, it appears that subjects are willing to pay for energy savings but not at the expense of certain conveniences and not without actual cost data for price/energy saving trade-offs. Overall, the evidence shows that consumers' existing knowledge in the area of energy and its relation to products and other costs is sadly inadequate, but the provision of objective information of the type explored in this study holds promise as a valuable decision making tool for consumers.

Conclusion

The public policy environment provides a unique opportunity for consumer researchers, if he or she is aware of factors that distinguish policy-oriented research from that done for other purposes. In particular, four "generalizations" may be useful for empirical researchers considering research in this area:

- Policy objectives are crucial in determining the research approach. In the short run, emphasis is necessarily placed on problems solving. But, in the long run, research addressing broader issues of CIP is necessary for a fuller understanding of consumer information use as well as for assessing the effects of information.
- Experiments show particular promise as a vehicle for both research approaches.
- For the policy-oriented researcher, the design is very much a function of the population under study and the problem to be addressed.
- Task environment, measures, and stimulus vehicles are particularly important because of the stress on external validity. Careful attention should be paid to all three in the interest of realism and to counter potential demand characteristics.

Anyone taking these seriously will find ample challenge in terms of research creativity, as well as having the opportunity to contribute something useful to society, our profession, or both.

References

- James R. Bettman, "Issues in Designing Consumer Information Environments," Journal of Consumer Research, 2(December 1975), 169-177.
- James R. Bettman and Pradeep Kakkar, "Effects of Information Presentation Format on Consumer Information Acquisition Strategies," Journal of Consumer Research, 3(March 1977), 233-240.
- Donald T. Campbell and Julian C. Stanley, Experimental and Quasi-Experimental Designs for Research, (Chicago: Rand McNally, 1963).
- George S. Day, "Assessing the Effects of Information Disclosure Requirements," Journal of Marketing, 40 (April 1976), 45-52.

R.M. Heeler and M.L. Ray, "Measures Validation in Marketing," Journal of Marketing Research, 9(November 1972), 361-370.

Michael J. Houston, "The Effect of Unit Pricing on Choices of Brand Size in Economic Shopping Behavior," Journal of Marketing, 36(July 1972), 51-54.

R. Bruce Hutton, "The Impact on the Processing of New Information for Durable Goods," unpublished Ph.D. dissertation, University of Florida, 1977.

Hans R. Isakson and Alex R. Maurizi, "The Consumer Economics of Unit Pricing," Journal of Marketing Research, 10(August 1973), 227-235.

Fred N. Kerlinger, Foundations of Behavioral Research, 2nd Ed., (New York: Holt, Rinehart, and Winston, Inc., 1973).

Sanford I. Labovitz, "Criteria for Selecting a Significance Level: A Note on the Sacredness of .05," in D.P. Forcese and S. Richer, eds., Stages of Social Research: Contemporary Perspectives, (New Jersey: Prentice-Hall, Inc., 1970).

B.J. McGuire and E.A. Vadelund, Voluntary Labeling Program for Household Appliances and Equipment to Effect Energy Consumption: Annual Report for Calendar Year 1974, Center for Consumer Technology, Institute for Applied Technology, National Bureau of Standards, Washington, D.C., February 1975.

Dennis L. McNeill, "An Investigation of Issues Surrounding the Disclosure of Energy Use Data on Major Home Appliances," unpublished Ph.D. dissertation, University of Florida, 1977.

J. Edward Russo, Gene Krieser, and Salley Miyashita, "An Effective Display of Unit Price Information," Journal of Marketing, 39(April 1975), 11-19.

Alan G. Sawyer, "Demand Artifacts in Laboratory Experiments in Consumer Research," Journal of Consumer Research, 1(March 1975), 20-30.

_____, The Energy Policy and Conservation Act, Public Law 94-163, 1975.

_____, The M.I.T. Report: Consumer Appliances, the Real Cost, RANN Document Center, National Science Foundation, Washington, D.C., 1974.

_____, The Productivity of Servicing Consumer Durable Products, The Center for Policy Alternatives, Massachusetts Institute of Technology (M.I.T.) with Charles Stark Draper Laboratory, Inc., Report No. CPA-74-4, 1974a.

William L. Wilkie, "Analysis of Effects of Information Load," Journal of Marketing Research, 11(November 1974).

William L. Wilkie, How Consumers Use Product Information. Report prepared for National Science Foundation, Grant No. GI-42037, Washington, D.C.: Government Printing Office, 1975a.

William L. Wilkie, "Consumer Information Processing Research: Product Labeling," Paper Series; Marketing Science Institute, 1975b.

William L. Wilkie, "New Perspectives for Consumer Information Processing Research," Communication Research, 2 (July 1975c), 216-231.

William L. Wilkie and David M. Gardner, "The Role of Marketing Research in Public Policy Decision Making," Journal of Marketing, 38(January 1974), 38-47.

William L. Wilkie and R. Bruce Hutton, "Life Cycle Cost: Impact on Consumer Response to New Product Information," in progress, 1977.

William L. Wilkie, R. Bruce Hutton, and Dennis L. McNeill, "Application of Consumer Information Processing Research to Public Policy," Southern Marketing Association Conference Proceedings, November 1975.

POTENTIAL CONTRIBUTIONS OF CONSUMER RESEARCH TO ANTITRUST DECISION MAKING*

Paul N. Bloom, University of Maryland

Abstract

A review is presented of several antitrust issues that might be resolved with the help of consumer research. Some research methodologies that could be used to examine these issues are also suggested. It is concluded that consumer research could play an important role in antitrust decision making.

Introduction

Numerous authors have written about the need for consumer research to help public policy decision making (Wilkie and Gardner, 1974; Rosch, 1975; Hunt, 1976). These authors have stimulated a growing volume of research studies designed to assist policy making in the consumer protection area (Day and Brandt, 1974; Jacoby, Speller, and Kohn, 1974a, 1974b) and other more "functional" areas (Richie and LaBreque, 1975) such as pollution control (Kinnear, Taylor, and Ahmed, 1974), transportation (Brown and Schary, 1977), and energy conservation (Reizenstein and Barnaby, 1977). However, one area of public policy decision making that has tended to be overlooked by consumer researcher is antitrust policy. Although several authors have recognized that consumer research could aid antitrust enforcement efforts (Wilkie and Gardner, 1974; Howard, 1977; Day and Shocker, 1977), very little research has been done with the specific goal of helping antitrust decision making. Moreover, rarely have consumer researchers considered whether there are antitrust implications in the findings of any of their studies.

In an effort to stimulate more research and thinking about antitrust on the part of consumer researchers, this paper contains a discussion of several areas where consumer research could contribute to antitrust decision making. The focus of the discussion is on how consumer research could help to resolve some of the less traditional issues that have arisen in current antitrust cases involving the four major ready-to-eat breakfast cereal manufacturers, the legal and medical professions, and ReaLemon. However, the discussion also briefly covers how consumer research could help to resolve more traditional antitrust questions involving subjects such as tying agreements, defining market boundaries, and parallel pricing.

It should be noted that the following is not intended to provide an exhaustive survey of antitrust issues that might be resolved with the help of consumer research; nor is it intended to provide an exhaustive review of previous consumer research findings that could contribute to the resolution of antitrust issues. The discussion is merely intended to suggest a few antitrust issues that are ripe for consumer research, and to point out some previous research efforts that might illustrate methods for resolving these issues.

Table 1 contains an outline of the types of antitrust issues or questions that consumer research could help to resolve. The body of the paper is organized to follow this outline. The concluding section of the paper contains a short discussion of some of the practical

*This research was supported in part by a grant from the Bureau of Business and Economic Research of the University of Maryland.

problems that could arise in trying to expand the role of consumer research in antitrust decision making.

TABLE 1

Selected Antitrust Issues that Consumer Research Might Help to Resolve

- I. Case Selection Issues
 - A. When is an industry or firm performing poorly?
 - B. Do the expected benefits from filing a suit or complaint exceed the expected litigation costs?
- II. Case Resolution Issues
 - A. Structural Characteristics Questions
 1. What is the relevant market?
 2. How high are product differentiation barriers to entry?
 3. Are there economies of scale in advertising?
 - B. Anticompetitive Practices Questions
 1. Are brand proliferation and product differentiation strategies exclusionary?
 2. Is intensive advertising exclusionary?
 3. Is advertising pro-competitive?
 4. Are shelf-space allocation programs exclusionary?
 5. When is the integrity of a firm's product damaged?
 6. When are two products of like grade and quality?
 7. When is parallel pricing justified?
- III. Case Remedy Selection Issues
 - A. What would be the effects of discontinuing certain anticompetitive practices?
 - B. Will royalty-free licensing of brand names and trademarks encourage more competition?
 - C. Will new information presentation formats encourage more competition?

Case Selection Issues

The procedures followed by the Antitrust Division of the Justice Department and the Federal Trade Commission in selecting antitrust cases are not well-defined or widely understood. Clearly, consideration is given to a host of factors in deciding whether to take action against a given industry or firm. Among the factors generally considered are: the structure, conduct, and performance of the suspect industry; the likelihood of winning the case; the amount of agency resources needed for litigation; and the political pressures favoring or

opposing a suit or complaint (Green 1972). There are at least two ways consumer research could help to make the case selection process more efficient and equitable.

First, consumer research could be used to develop improved measures of the "performance" of an industry or firm. At present, the antitrust enforcement agencies utilize performance measures suggested by industrial organization economists. These include measures of allocative efficiency (rates of return on equity), technical efficiency (production costs), and so on (Stern and Grabner, 1970). An additional performance measure that could be of value to the agencies is "consumer satisfaction" (or dissatisfaction). The techniques currently being developed by consumer researchers to measure consumer satisfaction, through monitoring consumer complaints and other procedures (Aiello, Czepiel, and Rosenberg, 1977; Handy, 1977), could be used by the agencies to detect which industries or firms are producing relatively undesirable outcomes through their methods of competition.

Second, consumer research could be used to estimate the social benefits of winning alternative antitrust cases before a complaint or suit is filed. These estimates could help the enforcement agencies decide whether it would be a wise investment of public funds to litigate a particular case (Hunt, 1976). The kinds of research that would help to make these estimates would be essentially the same as the kinds of research needed to determine the effects of certain questionable business practices or the effects of certain proposed antitrust remedies. These research areas are discussed in the following sections.

Case Resolution Issues

Once an antitrust suit or complaint has been filed, consumer research could help to resolve a wide range of issues. Testimony and evidence drawn from consumer research studies could help the courts, administrative law judges, and FTC obtain (1) a more accurate picture of the structural characteristics of a given industry and (2) a better estimate of the anticompetitive effects of certain business practices. Potential research contributions in both of these areas are discussed at length below.

Structural Characteristics Questions

One of the most basic questions that must be answered in any antitrust case is: What is the relevant market? This question must be answered before concentration ratios can be computed, injured competitors can be identified, or other important judgments can be made. Historically, the courts have determined relevant markets by looking at evidence such as cross-elasticity of demand figures, statements and memos from employees concerning who they see as their competitors, and descriptions of manufacturing processes (Day, Massy, and Shocker, 1977). But several observers have recommended that more attention should be paid to consumer perceptions of competing products in determining relevant markets (Stern and Grabner, 1970; Day and Shocker, 1977).

In a recent paper, Day and Shocker (1977) review an assortment of consumer research approaches that could be used to obtain consumer perceptions of market boundaries. They recommend the concurrent use of several of these approaches in defining any given market. Suggested approaches include (1) the study of brand-switching rates (to see which brands are used as substitutes for one another), (2) the use of decision sequence analysis to obtain protocols of brand choice processes (Bettman, 1971, 1974) (to see which other brands are considered when making a choice), (3) the

use of direct consumer judgments of substitutability, perhaps employing techniques such as the "dollar metric" approach of Pessemier et al. (1971) (to see how much money consumers would want before substituting one brand for another), and (4) the use of perceptual mapping techniques. Etgar and Jain (1976) demonstrated how this last approach might be used when they employed multi-dimensional scaling and cluster analysis techniques to identify market boundaries for beverages.

In addition to helping to identify the competitors in a given market, consumer research could be useful in determining the extent of product differentiation in the market or the height of product differentiation barriers to entry. Barriers of this type are said to exist when brand loyalty makes it difficult for new brands to attract buyers without spending large sums on advertising and promotion (Bain, 1956; Comanor and Wilson, 1974). Whether these barriers exist has been an issue in several antitrust cases--the most notable being the cereal case (Root, 1972).

Research on the degree of brand loyalty in several different industries could shed light on the height of product differentiation barriers in any one of those industries. The existence of relatively strong brand loyalties in one industry would indicate that differentiation was great and that entry barriers were relatively high in that industry. Data on brand loyalty could be obtained from consumer surveys, panel diaries, and experimental investigations. The last approach might make use of procedures similar to those used by Ehrenberg and Charlton (1973, 1976) to assess how difficult it is to use various marketing tactics to overcome brand loyalty. They have conducted several longitudinal, quasi-field experiments on repeat-buying, using door-to-door salespersons to administer treatments and obtain measures.

Another structural characteristic of a market that could be important in antitrust cases, and that could be studied through consumer research, is the size of economies of scale in advertising. Such economies could serve to bar entry by smaller competitors by putting them at a cost disadvantage. The possibility that such economies could emerge was cited by the Supreme Court as a reason for voiding the merger of Procter and Gamble and Clorox (Fusilier and Darnell, 1971).

Economies of scale in advertising can emerge if quantity discounts are offered for media purchases or if some minimum number of advertising messages must be received by consumers before the advertising becomes effective--i.e., a threshold level of advertising expenditures must be reached, above which increasing returns to advertising can be obtained (Comanor and Wilson, 1974). Experimental studies with consumers could conceivably be used to gain insights into whether economies of scale in advertising can emerge because of a threshold effect. Some of the research that has been done to examine how consumers react to repeated persuasive communications over time (Ray and Sawyer, 1971; Mitchell and Olson, 1977) could possibly be extended to study the economies of scale issue.

Anticompetitive Practices Questions

The focus of the discussion in this section is on business practices that have been accused of being exclusionary or restrictive in recent antitrust cases. While the use of deceptive or misleading practices has also been an issue in several recent antitrust cases, the need for research on the competitive effects of these practices will not be elaborated upon here. Consumer researchers have already recognized the need to study the effects of deceptive and misleading practices (Gardner, 1975; Russo, 1976), although previous

research has tended to be most concerned with how these practices hurt consumers as opposed to competition.

Four practices that have been accused of having exclusionary effects in the cereal case are: brand proliferation, product differentiation, intensive advertising, and shelf-space allocation. In a trial currently in progress, the FTC staff is presenting a complex set of interwoven arguments to convince an administrative law judge that the use of these four practices has allowed the four largest manufacturers of ready-to-eat breakfast cereals to exclude entry and maintain a "shared monopoly" (FTC Complaint Counsel, 1976). The arguments of the FTC staff are summarized (and somewhat simplified) below, accompanied by a discussion of how consumer research might be used to test the validity of these arguments.

The FTC staff is first arguing that the major cereal manufacturers (Kellogg, General Mills, General Foods, and Quaker) have essentially used brand proliferation and product differentiation to monopolize every profitable "position" in the "product space" of the cereal market. Describing a multi-attribute product space of the type presented by Lancaster (1975), the FTC staff claims that:

1. All profitable positions in the product space are currently occupied by brands of the major manufacturers.
2. With so many brands competing to let consumers know where they stand in the space, the costs to a new brand of getting the attention of consumers and establishing itself in a preferred location in the space are very high.
3. With so many brands already established in preferred locations in the space, it is virtually impossible for one brand to create enough preference for itself to obtain a share of the cereal market large enough to take advantage of economies of scale in production. This means that to attain competitive production costs a new firm will need several successful brands established in multiple locations across the product space. The promotional costs or lower prices needed to establish several brands in preferred locations in the space serve to deter entry.
4. The entry of a new firm in an unoccupied, profitable position would only attract the introduction of new, established-firm brands in that position.
5. The proliferation of numerous differentiated brands confuses consumers and makes it difficult for them to recognize the physical similarities that exist among brands. This makes it more difficult for new brands to become established in preferred locations in the space--particularly private labels that are meant to be seen as copies of existing brands.

Testing the validity of the above arguments would surely be a challenging task. Nevertheless, there are several consumer research approaches that might prove helpful. For instance, perceptual mapping or multidimensional scaling techniques could be used to determine whether, from the perspective of a representative sample of consumers, all desirable locations are occupied in the product space for the cereal market. If "ideal points" were found that are not being served by existing brands, then the FTC staff would have a weaker case. In addition, perceptual mapping techniques could be used to test how confused consumers are about the attributes of different brands. If the perceptual maps of cereal brands obtained from consumers did not differ greatly from maps drawn using objectively-measured characteristics, the argument about consumer confusion could be questioned.

Some insights into the validity of the above arguments could also be obtained by conducting more general (less case-specific) experimental studies with several classes of products. Experiments could test whether increasing the number of alternative brands in a product category produces effects such as:

1. Increased confusion for consumers when processing information about the brands. The research being done on "information overload" (Jacoby, Speller, and Kohn, 1974a, 1974b) suggests that this does, in fact, happen, but more research on this matter is definitely needed.
2. Increased difficulty in using persuasive communications to alter the cognitive structures of consumers and establish a brand in a preferred location in the product space. The research that is being done on the use of persuasive communications to alter cognitive structures or perceptual maps (Olson and Dover, 1976; Lutz, 1975; Moinpoir et al., 1976) could be extended to test whether having more brands means that more communications are needed to change beliefs, salient attributes, ideal points, and so forth.

If these experiments found that increasing the number of brands did not increase confusion or the difficulty of "positioning" brands in desirable locations, then the arguments against brand proliferation and product differentiation would be weakened.

Finally, studies about the variety-seeking or novelty-seeking behavior of consumers (Venkatesan, 1973; Raju, 1976) could also prove helpful in evaluating the arguments about brand proliferation and product differentiation. If it were found that consumers obtain more satisfaction over time when large numbers of differentiated brands are made available--because they tend to try out more brands--then a case could be made for claiming that brand proliferation and product differentiation produce important social benefits.

Intensive advertising is another practice being attacked in the cereal case. The FTC staff is arguing that this practice excludes entry to the industry in the following ways:

1. It supports the brand proliferation and product differentiation activities.
2. It allows existing large firms to take advantage of economies of scale in advertising which new small firms cannot obtain.
3. It creates brand loyalty which makes it difficult for new brands to attract customers--i.e., product differentiation barriers to entry are created.

The latter two charges are arguments that have frequently been offered by economists who view advertising as being anticompetitive (Bain, 1956; Comanor and Wilson, 1974).

As discussed earlier, consumer research could help to establish whether economies of scale in advertising or strong brand loyalties exist in a given market. To be sure, the discovery of advertising scale economies would indicate that the use of intensive advertising was probably having exclusionary effects. However, the discovery of strong brand loyalties would not necessarily suggest anything about the effects of intensive advertising. Loyalty could be created by many factors other than advertising. In fact, attempting to isolate the role that advertising plays in creating consumer brand loyalty represents an exceedingly difficult research problem.

Although it might be possible to use experimental procedures such as those used by Ehrenberg and Charlton (1973, 1976) to determine whether intensive advertising can create strong brand loyalties, it would probably be more fruitful to use consumer research to evaluate, in a general way, several arguments that will probably be offered in the cereal case in defense of the use of intensive advertising. These pro-advertising arguments have been offered by economists such as Brozen (1974), Ferguson (1974), and Nelson (1974, 1975) to counter claims that advertising is anticompetitive. If these pro-advertising arguments are shown to have some validity, then the case against intensive advertising would be weakened (or vice versa). In addition, evaluations of these pro-advertising arguments would probably prove helpful in reaching decisions about the appropriateness of permitting advertising in the professions. Interestingly enough, the FTC and Antitrust Division have essentially adopted these pro-advertising arguments in their cases against lawyers, doctors, pharmacists, and optometrists.

Economists such as Brozen and Nelson have basically argued that advertising is pro-competitive because it makes demand curves more elastic. It does this, in the case of goods which can be evaluated without actually trying them (Nelson's "search" goods), by providing inexpensive information to consumers and making it cheaper for them to comparison shop. With more comparison shopping being done as a result of advertising, consumers will react more sharply to price changes and thus encourage more price competitiveness on the part of sellers.

In the case of goods which cannot be evaluated without actually trying them (Nelson's "experience" goods), it is argued that advertising still makes demand curves more elastic, but for somewhat different reasons. Advertising informs consumers about which brands are the "best buys" (i.e., the lowest price per unit of utility). It does this, according to Nelson (1974, 1975), because the best buys will be the most heavily advertised brands. Heavier advertising would be done for the best buys basically because it does not pay to advertise a bad product and because most firms set their advertising budgets on a percentage of sales basis. With consumers knowing which brands are best, they can limit their sampling or trial of brands to a relatively small, homogeneous group and not be forced to consider a broad range of heterogeneous brands. Consumers will therefore react more sharply to price changes of brands in their "evoked set" when advertising exists than they would without advertising--since they will be more likely to notice price changes when less brands are considered for trial. Thus, advertising should produce more price competitiveness among sellers.

The pro-advertising authors have generally turned to studies using aggregate archival data to find support for their arguments. Two supportive studies that have frequently been cited are the works which found prices of eyeglasses (Benham, 1972) and prescription drugs (Cady, 1975) to be lower in states that permitted advertising of these goods. However, experimental studies obtaining data directly from consumers might also prove useful in evaluating the pro-advertising arguments. Experiments could be set up to study how the information-acquisition strategies, brand choices, and post-choice satisfaction of consumers are affected by varying (1) the amount of advertising available in a product category (none, moderate for a few brands, intensive for a few brands) and (2) the type of product category ("search" vs. "experience" goods or tangible vs. intangible goods). These experiments could attempt to measure the costs of search incurred by subjects under various experimental conditions by charging the subjects fees for information obtained (the fee for using

advertising information would be set relatively low). Information display boards (Berning and Jacoby, 1974; Jacoby et al., 1976), verbal protocols (Bettman and Jacoby, 1976; Bettman, 1971, 1974), and eye-movement measures (Russo and Doshier, 1976) could all be used to obtain a better understanding of how advertising information is processed. The pro-advertising arguments would tend to be supported if the treatment groups which were exposed to brand advertising incurred lower search costs, examined a smaller set of advertised brands, chose brands which were better buys, and obtained more satisfaction.

Shelf-space allocation is the final practice that is being challenged by the FTC staff in the cereal case. The four major manufacturers have been charged with tacitly agreeing to persuade retailers to utilize a single shelf-space plan (Kellogg's) which: (1) gives their brands "better" locations at the center of the aisle and relegates other brands to "poorer" locations at the ends of the aisle, (2) has the brands of each manufacturer displayed in a different grouping, making it more difficult for consumers to compare similar brands (e.g., corn flakes) of different manufacturers, and also making it more likely that consumers will select second brands made by the same company which manufactured their first selection, and (3) has their brands "billboarded" in the displays (several facings are put next to one another), getting them more attention and impulse purchases. Clearly, experimental consumer research studies in either simulated shopping environments or real-world supermarkets, along the lines of work done by Cox (1970) and Curhan (1972), could be used to test the validity of these arguments.

While the previous discussion has focused on business practices that have been questioned in the cereal case and the cases involving the professions, there are many other practices that have been challenged in antitrust cases that could be evaluated using consumer research approaches. For example, surveys of consumers could probably provide guidance in cases involving tying agreements and dealer restrictions. Surveys of consumer attitudes and preferences could help to establish whether a manufacturer or franchisor could use the defense that the integrity of its product has been damaged when these restrictive practices have not been used. Similarly, surveys could be used in price discrimination cases to help determine whether consumers see two products as being of "like grade and quality" (Day, Massy, and Shocker, 1977). In addition, experiments with consumers could probably provide guidance in cases involving parallel pricing. Experiments on the sensitivity of consumers to price cuts in a certain market, or on consumer perceptions of a price-quality relationship, could help to establish whether the sellers in that market could defend parallel pricing behavior by arguing that consumers do not respond to lower prices.

Case Remedy Selection Issues

Since many antitrust cases are concluded by voiding a merger or acquisition or by requiring firms to discontinue the use of an anticompetitive practice, the type of research discussed in the previous sections would obviously be useful in selecting case remedies. However, other forms of consumer research might prove helpful for locating effective case remedies that fall somewhere in between the restructuring of industries (which rarely is done) and the ordering of firms to discontinue the use of a practice (which may not restore competition to its former, pre-violation state). An example of this type of remedy is one currently being sought by the FTC staff in both the ReaLemon case and the cereal case: the royalty-free licensing of brand names and trademarks to new firms.

The FTC staff believes that royalty-free licensing will encourage new manufacturers to produce and market popular brands (e.g., ReaLemon reconstituted lemon juice, Rice Krispies cereal, Cheerios cereal), while at the same time discourage existing manufacturers from using practices such as brand proliferation, product differentiation, and intensive advertising. New manufacturers could obtain a valuable brand name at no cost, while existing manufacturers would no longer have an incentive to spend large amounts marketing a brand because they might only be helping a competitor by doing so. The result would presumably be more price competition among groups of homogeneous brands.

An implicit and crucial assumption that seems to exist in the arguments for royalty-free licensing is that consumers will react in a similar manner to Safeway's ReaLemon or Kroger's Rice Krispies as they have to Borden's ReaLemon and Kellogg's Rice Krispies. Such an assumption could probably be tested using experimental consumer research approaches. If it were found, for example, that a company name can serve to differentiate a product in the eyes of consumers almost as well as a brand name, then the new remedy would most likely prove to be ineffective for encouraging more competitive activity.

It might also prove useful to turn to past consumer research for ideas for additional antitrust remedies. For example, it has been found that information presentation format can have a significant effect on the way consumers acquire and process information (Bettman and Kakkar, 1977; Russo et al., 1975)--i.e., that certain displays will lead consumers to process information by brand and others will encourage processing by attribute. It has also been found (in a study of cereal brands) that processing by attribute is more characteristic of consumers who exhibit low brand loyalty (Jacoby et al., 1976). These findings suggest that an effective remedy in the cereal case might be one which required supermarkets to display cereal brands in a way that encouraged attribute processing of information--perhaps by arranging brands according to types of ingredients or by erecting signs which compare brands on certain attributes (with one attribute per sign). More attribute processing by cereal customers might conceivably lead to less brand loyalty, more comparative shopping, and more competition among cereal manufacturers. Of course, experimental tests of this remedy would be needed, since low brand loyalty might lead to attribute processing without the reverse being true.

Conclusion

Consumer research can clearly play an important role in the enforcement of the antitrust laws. There are many case-specific studies that could be done to help resolve individual antitrust cases. Moreover, there are a number of more general studies that could be done that might provide insights applicable to several cases. In particular, studies of the effects of brand proliferation, product differentiation, and intensive advertising could provide valuable guidance in attempts to resolve either present cases, like those involving the cereal industry and the professions, or potential future cases involving the manufacturers of automobiles, detergents, or other highly differentiated consumer goods.

Whether antitrust lawyers will be willing to submit the results of consumer research studies as courtroom evidence is another issue. As several authors have pointed out (Hunt, 1976; Gardner, 1974), there are substantial problems associated with submitting consumer research in court. It is also unclear whether a willingness to use consumer research results will be exhibited by those involved with selecting new cases or devising new remedies. These lawyers and economists generally have a

weak understanding of the measurement techniques and research approaches used by consumer researchers. Nevertheless, given the growing tendency of FTC officials to utilize consumer research results in consumer protection decision making (Brandt and Preston, 1977), the prospects seem reasonably good for also seeing some increase in the use of consumer research in the antitrust area. Word-of-mouth advertising for consumer research by personnel in the FTC's Bureau of Consumer Protection seems likely to spark the interest of antitrust enforcers at the FTC's Bureau of Competition and Bureau of Economics, as well as at the Justice Department's Antitrust Division. Thus, if consumer researchers can begin to supply research findings that antitrust people find relevant and timely, the use of consumer research in antitrust decision making could expand rapidly.

References

- Aiello, Albert, John A. Czepiel, and Larry J. Rosenberg. "Scaling the Heights of Consumer Satisfaction: An Evaluation of Alternative Measures." Paper presented at the Research Symposium on Consumer Satisfaction/Dissatisfaction and Complaining Behavior, Indiana University, April 21 and 22, 1977.
- Bain, Joe S. Barriers to New Competition. Cambridge, Mass.: Harvard University Press, 1956.
- Benham, Lee. "The Effect of Advertising on the Price of Eyeglasses," Journal of Law and Economics, 15 (October 1972), 337-352.
- Berning, Carol A. Kohn and Jacob Jacoby. "Patterns of Information Acquisition in New Product Purchases," Journal of Consumer Research, 1 (September 1974), 18-22.
- Bettman, James R. "The Structure of Consumer Choice Processes," Journal of Marketing Research, 8 (November 1971), 465-471.
- _____. "A Threshold Model of Attribute Satisfaction Decisions," Journal of Consumer Research, 1 (September 1974), 30-35.
- Bettman, James R. and Jacob Jacoby. "Patterns of Processing in Consumer Information Acquisition," in B. B. Anderson (Ed.), Advances in Consumer Research, Volume 3. Chicago: Association for Consumer Research, 1976, 315-320.
- Bettman, James R. and Pradeep Kakkar. "Effects of Information Presentation Format on Consumer Information Acquisition Strategies," Journal of Consumer Research, 3 (March 1977), 233-240.
- Brandt, Michael T. and Ivan L. Preston. "The Federal Trade Commission's Use of Evidence to Determine Deception," Journal of Marketing, 41 (January 1977), 54-62.
- Brown, Daniel J. and Philip B. Schary. "Consumer Participation in Transportation Planning: The Elderly, the Poor, and Special Interests," in William D. Perreault, Jr. (Ed.), Advances in Consumer Research, Volume 4. Chicago: Association for Consumer Research, 1977, 138-141.
- Brozen, Yale. "Is Advertising a Barrier to Entry?" in Yale Brozen (ed.), Advertising and Society. New York: New York University Press, 1974, 79-110.
- Cady, John F. Drugs on the Market. Lexington, Mass.: Lexington Books, 1975.

- Charlton, P. and A. S. C. Ehrenberg. "An Experiment in Brand Choice," Journal of Marketing Research, 13 (May 1976), 152-160.
- Comanor, William S. and Thomas Wilson. Advertising and Market Power. Cambridge, Mass.: Harvard University Press, 1974.
- Cox, Keith K. "The Effect of Shelf Space Upon Sales of Branded Products," Journal of Marketing Research, 7 (February 1970), 55-58.
- Curhan, Ronald C., "The Relationship Between Shelf Space and Unit Sales in Supermarkets," Journal of Marketing Research, 9 (November 1972), 406-412.
- Day, George S. and William K. Brandt. "Consumer Research and the Evaluation of Information Disclosure Requirements: The Case of Truth in Lending," Journal of Consumer Research, 1 (June 1974), 21-32.
- Day, George S., William F. Massy, and Allan D. Shocker. "The Public Policy Context of the Relevant Market Question," working paper, University of Toronto, 1977.
- Day, George S. and Allan D. Shocker. "Identifying Competitive Product Marketing Boundaries: Strategic and Analytic Issues," Working Paper No. 156, Graduate School of Business, University of Pittsburgh, 1977.
- Ehrenberg, A. S. C. and P. Charlton. "An Analysis of Simulated Brand--Choice," Journal of Advertising Research, 13 (February 1973), 21-23.
- Etgar, Michael and Arun K. Jain. "How to Improve Antitrust Policies with Market Research Tools," in Edward M. Mazze (Ed.), 1975 Combined Proceedings. Chicago: American Marketing Association, 1976, 72-75.
- Ferguson, James M. Advertising and Competition: Theory, Measurement, Fact. Cambridge, Mass.: Ballinger Publishing Company, 1974.
- FTC Complaint Counsel, Trial Brief--In the Matter of Kellogg Company et al., Docket No. 8883, Federal Trade Commission, 1976.
- Fusilier, H. Lee and Jerome C. Darnell. Competition and Public Policy: Cases in Antitrust. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1971.
- Gardner, David M. "Dynamic Homeostasis: Behavioral Research and the FTC," in Scott Ward and Peter Wright (Eds.), Advances in Consumer Research, Volume 1. Chicago: Association for Consumer Research, 1974, 108-113.
- Gardner, David M. "Deception in Advertising: A Conceptual Approach," Journal of Marketing, 39 (January 1975), 40-46.
- Green, Mark J. The Closed Enterprise System. New York: Grossman Publishers, 1972.
- Handy, Charles R. "Indexes of Consumer Satisfaction With Food Products: 1974 and 1976 Survey Results." Paper presented at the Research Symposium on Consumer Satisfaction/Dissatisfaction and Complaining Behavior, Indiana University, April 21 and 22, 1977.
- Howard, John A. Consumer Behavior: Application of Theory. New York: McGraw-Hill Book Company, 1977.
- Hunt, H. Keith. "What About Marketing Research and the Law at the Federal Trade Commission," in Edward M. Mazze (Ed.), 1975 Combined Proceedings. Chicago: American Marketing Association, 1976, 85-89.
- Jacoby, Jacob, Robert W. Chestnut, Karl C. Weigl and William Fisher. "Pre-purchase Information Acquisition: Description of a Process Methodology, Research Paradigm, and Pilot Investigation," in B. B. Anderson (Ed.), Advances in Consumer Research, Volume 3. Chicago: Association for Consumer Research, 1976, 306-314.
- Jacoby, Jacob, Donald E. Speller, and Carol A. Kohn. "Brand Choice Behavior as a Function of Information Load," Journal of Marketing Research, 11 (February 1974a), 63-69.
- _____. "Brand Choice Behavior as a Function of Information Load: Replication and Extension," Journal of Consumer Research, 1 (June 1974b), 33-42.
- Kinnear, Thomas C., James R. Taylor, and Sadrudin A. Ahmed. "Ecologically Concerned Consumers: Who Are They?" Journal of Marketing, 38 (April 1974), 20-24.
- Lancaster, Kelvin. "Socially Optimal Product Differentiation," American Economic Review, 65 (September 1975), 567-585.
- Lutz, Richard J. "Changing Brand Attitudes Through Modification of Cognitive Structure," Journal of Consumer Research, 1 (March 1975), 49-59.
- Mitchell, Andrew A. and Jerry C. Olson. "Cognitive Effects of Advertising Repetition" in William D. Perreault, Jr. (Ed.), Advances in Consumer Research, Volume 4. Chicago: Association for Consumer Research, 1977, 213-220.
- Moinpoir, Reza, James M. McCullough, and Douglas MacLachlan. "Time Changes in Perception: A Longitudinal Application of Multidimensional Scaling," Journal of Marketing Research, 13 (August 1976), 245-253.
- Nelson, Phillip. "Advertising as Information," Journal of Political Economy, 81 (July/August 1974), 729-754.
- _____. "The Economic Consequences of Advertising," Journal of Business, 48 (April 1975), 213-241.
- Olson, Jerry C. and Philip Dover. "Effects of Expectation Creation and Disconfirmation on Belief Elements of Cognitive Structure," in B. B. Anderson (Ed.), Advances in Consumer Research, Volume 3. Chicago: Association for Consumer Research, 1976, 168-175.
- Pessemier, Edgar A., Philip Burger, Richard Teach, and Douglas Tigert. "Using Laboratory Brand Preference Scales to Predict Consumer Brand Purchases," Management Science, 17 (February 1971), 371-385.
- Raju, P. S. "An Empirical Investigation of Variation Seeking in Purchase Behavior: A Complexity Theory Approach," Unpublished Ph.D. dissertation, University of Illinois, 1976.
- Ray, Michael L. and Alan G. Sawyer. "A Laboratory Technique for Estimating the Repetition Function for Advertising Media Models," Journal of Marketing Research, 8 (February 1971), 20-29.
- Reizenstein, Richard C. and David J. Barnaby. "The Consumer and the Energy Shortage: A Post-Embargo Assessment," in William D. Perreault, Jr. (Ed.), Advances in Consumer Research, Volume 4. Chicago: Association for Consumer Research, 1977, 308-314.
- Richie, J. R. Brent and Roger J. La Breque. "Marketing Research and Public Policy: A Functional Perspective," Journal of Marketing, 39 (July 1975), 12-19.
- Root, H. Paul. "Should Product Differentiation Be Restricted?" Journal of Marketing, 36 (July 1972), 3-9.

Rosch, J. Thomas. "Marketing Research and the Legal Requirements of Advertising," Journal of Marketing, 39 (July 1975), 69-72.

Russo, J. Edward. "An Empirically-Based Method for Identifying Misleading Advertising," working paper, Graduate School of Business, University of Chicago, 1976.

Russo, J. Edward and Barbara A. Doshier. "An Information Processing Analysis of Binary Choice," working paper, Grad. Sch. of Bus., Univ. of Chicago, 1976.

Russo, J. Edward, Gene Krieser, and Sally Miyashita. "An Effective Display of Unit Price Information." Journal of Marketing, 39 (April 1975), 11-19.

Stern, Louis W. and John R. Grabner, Competition in the Marketplace. Glenview, Illinois: Scott, Foresman and Company, 1970.

Venkatesan, M. "Cognitive Consistency and Novelty Seeking," in Scott Ward and Thomas S. Robertson (Eds.), Consumer Behavior: Theoretical Sources. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1973, 354-384.

Wilkie, William L. and David M. Gardner. "The Role of Marketing Research in Public Policy Decision Making," Journal of Marketing, 38 (January 1974), 38-47.

Debra L. Scammon, University of Utah

Abstract

Public policy is concerned with protecting the consumer from false, deceptive, and/or misleading brand information. This concern has led to a situation where a manufacturer may be required to disclose some brand information which may be discrepant with the consumers' currently held beliefs about the brand (i.e., corrective advertising, affirmative disclosure, ad substantiation). Little is known about the effects of such expectancy-discrepant information on consumers. This study investigates three areas of potential impact of expectancy-discrepant information on consumers--consumers' brand image/comprehension, consumers' brand perception in relation to an "ideal" brand, and consumers' psychological feelings about a brand selection decision. Two information disclosure formats are studied.

Introduction

Recently there has been extensive discussion, both in the marketing and social issues literature and in governmental hearings and meetings, regarding the appropriate role of policy-makers in consumer protection. This discussion has centered around determining the priority of issues appropriate to consumer protection programs. At the heart of much of this discussion lies the problem of availability of purchase-relevant information to the consumer and his/her use of this information in making purchase decisions. The discussion to follow focuses on one particularly crucial case of information availability, the case involving the disclosure of brand information contrary to or discrepant with the consumer's preconceived expectations about a brand. The effectiveness of two disclosure formats is compared.

Public Policy Interest

The concern of policy-makers with the quality of the consumer "information environment"² is not so much with how consumers interpret and use the information made available to them but rather whether or not purchase-relevant information is available to consumers if they want to use it in their brand choice decisions. As a result, information available about a brand may or may not be relevant to a particular consumer's decision to buy that brand. In addition, this information may or may not paint an accurate and realistic picture of the brand and its characteristics, i.e., the information may be false, deceptive, and/or non-existent. Upon search for and consideration of brand information, consumers may find

themselves in any of four situations:

1. They find they have no prior knowledge about the brand and thus no prior expectations about its contents, quality, performance, etc.
2. They find they have accurate and sufficient knowledge and expectations about the brand and need no further information to confirm their beliefs.
3. They find they have roughly accurate knowledge and expectations about the brand, but desire further confirmation to increase their confidence in their brand evaluations.
4. They find they have inaccurate knowledge and information about the brand and need new information to point out this discrepancy and convey accurate facts about the brand.

The second situation, where consumers have accurate and sufficient product information, may be the ideal from a consumer protection point of view. However, advertisers often find themselves facing consumers in either the third situation for established products where consumers have accurate but insufficient information to make a decision or the first situation for new products where consumers have insufficient information to make a decision. Thus, in some cases, for their own best interests, commercial sources can be expected to "fulfill" consumers' information needs. The issue of strict availability of information to consumers is obviously important then in the first and third cases above. The Nutrition Advertising Rule (39 Fed. Reg. 39482, 1974) and other affirmative disclosure actions have been proposed to assure the availability to the consumer of purchase-relevant facts when insufficient information or no data have previously been available to consumers.

However, it is in the fourth situation, where misinformation abounds, that public policy should place its regulatory emphasis. To correct erroneous product conceptions necessarily involves the disclosure of new information. For those consumers who have formed their product beliefs and/or brand comprehension on earlier inaccurate information, this new information is likely to be discrepant with their expectations about the brand. In this situation, the minimization of consumer deception and/or consumer biases founded on incorrect information can come from the discovery of the most effective way of correcting erroneous perceptions, images and/or product knowledge.

Effects on Consumers of Disconfirmation of Expectations

Studying the effects of disconfirmed expectations on consumers' product evaluations has been of interest from a marketing management point of view in attempting to maximize customer satisfaction with a product. Past research has used four theoretical models to predict the effects on product evaluations and customer satisfaction of discrepancies between expectations and actual or objective product performance:

1. Cognitive dissonance theory would suggest any discrepancy between actual and expected performance would be minimized by the consumer by adjustment of his perception of the product's performance to be more in line with his

¹The author would like to acknowledge the guidance of Harold Kassarjian and James Bettman, UCLA who served as dissertation committee chairman for the experimental study from which this paper was excerpted. Appreciation is also expressed to Kenneth Miller, University of Utah whose comments on earlier drafts of this paper were invaluable.

²The term "information environment" is used here to denote the entire array of product-related data available to the consumer. The term is adapted from James R. Bettman, "Issues in Designing Consumer Information Environments," *Journal of Consumer Research*, Vol. 2 (December 1975), pp. 169-177.

expectations of its performance, thus resulting in satisfaction.

2. Contrast theory suggests the consumer will magnify any discrepancy in performance between expectation and reality and will experience dissatisfaction with the actual performance.

3. Generalized negativity theory suggests that any discrepancy between expectation and reality will result in a negative evaluation of the product and a dissatisfied consumer.

4. Assimilation-contrast theory suggests that if the disparity between expectation and reality is small the consumer will assimilate the difference by perceiving the product's performance in line with his expectations. However, if the disparity is large the consumer will contrast or magnify the perceived disparity. Thus, a discrepancy could result in either a satisfied or a dissatisfied consumer.

Studies using these four theoretical models as bases for investigation have concluded that brand claims should not be overly exaggerated nor overly understated but rather, should be consistent with the actual attributes, quality, and performance of the brand. This honesty about the brand appears to create a high level of customer satisfaction with the brand once it is "experienced" (Anderson, 1973; Cardozo, 1965; Olshavsky and Miller, 1972).

An equally important question is how exposure to expectancy-discrepant information (other than experiential information) effects consumers' brand evaluations and subsequent satisfaction with the brand. Consumer protection legislation may involve requiring advertisers to give consumers information which is discrepant with consumers' current beliefs and expectations (especially through the use of remedies such as corrective advertising). In fact, it could be argued that there is little need for policy makers to be concerned with giving consumers information consistent with their current knowledge and beliefs if these are accurate.

When consumers have been given false or deceptive brand information, public policy may require the dissemination of information to correct any misconceptions based on the previously inaccurate information (i.e., corrective advertising). In effect, policy may require that consumers be given information disconfirming the expectations about a brand created through previous communication of false or deceptive information. Of interest then are the effects on consumers' brand evaluations of information deliberately designed to disconfirm their expectations. Exposure to such information could cause several quite different effects. Consumers could:

1. Believe the new information and change their attitude toward or perception of the brand.
2. Not believe the new information, mentally reject it, and continue believing what they had previously believed.
3. Not believe the new information but be curious about the discrepancy and seek further information to clarify the confusing situation.
4. Become confused and uncomfortable, not knowing what to believe (Maloney, 1962).

These possible reactions differ in their desirability from the point of view of the public policy objective underlying the information disclosure requirement. Thus, they have very different implications for public policy formulation. It is therefore critical to determine which reaction is the most likely response and whether the format of the disclosure influences the response.

A Test of the Effects of Two Formats for the Disclosure of Expectancy-Discrepant Information

The Experimental Design

Past research on information processing suggests that the degree to which new information is perceived as discrepant with currently held beliefs might be influenced both by the amount of new (discrepant) information presented and by the format of that information, i.e., its complexity/simplicity (Jacoby, et.al., 1974a,b; Russo, 1975).

As part of a more extensive investigation of the effects of amount and format of information on consumers' product evaluations, an experiment was designed in which respondents were assigned to groups receiving different amounts of purchase-relevant information in different formats (Scammon, 1976). The proposed Trade Regulation Rule on Nutrition Advertising (39 Fed. Reg. 39482, 1974) provided an appropriate framework for this analysis since its guidelines suggested different disclosure formats (%Recommended Daily Allowance or adjectival descriptions of the nutrient content) and several amounts of information (zero, four, or eight nutrients) which advertisers might be required to use when presenting nutrition claims about their products.

It was necessary to be able to introduce information discrepant with preconceived notions about the test brands. To do this, two brands of peanut butter (Skippy Peanut Butter and Koogole Peanut Butter Surprise) were selected. Skippy was significantly more well known than was Koogole and Skippy had a well established image in the minds of consumers. Skippy was rated significantly more well-known, more wholesome, and higher on nutrition value than was Koogole by a control group receiving no nutrition information on either brand (see Table 1).

TABLE 1
Brand Image Profiles
Mean Ratings of Brands on Selected Attributes¹

<u>Attribute</u>	<u>Skippy</u>	<u>Koogole</u>
Wholesome/unhealthy	1.98 ²	3.36
High on Nutrition/low on nutrition	1.98 ²	3.40
Well Known/unknown	1.38 ²	3.24

¹The control group was used as a surrogate measure of pre-exposure image of the test brands. Differences in means between the two brands significant beyond the .01 level of significance.

²5-point scale with 1.0 indicating the most positive response.

To maximize the possibility of image change, during the experiment Skippy was portrayed as the "nutritionally inferior" brand, an image discrepant with the one revealed by the control group. The information on Koogole was left unmodified so it was not discrepant with Koogole's current image.

The nutrition information was presented to respondents in 30-second color commercials for the two brands which had been modified by superimposing replicas of nutrition labels with the appropriate formats (% RDA or adjectival descriptions) and amounts (zero, four, or eight nutrients) of nutrition information over the last six seconds of the commercials as proscribed in the Nutrition Advertising Rule (see Figure 1).

FIGURE 1
Sample Disclosure Formats

KOOGLE PEANUT BUTTER SURPRISE
NUTRITION INFORMATION
PER SERVING

CALORIES = 180

PERCENTAGE OF U.S. RECOMMENDED DAILY ALLOWANCES (U.S. RDA)			
PROTEIN	35	CALCIUM	10
NIACIN	30	RIBOFLAVIN	10
THIAMINE	20	VITAMIN A	0
IRON	20	VITAMIN C	0

1a. Sample disclosure format using percentage of Recommended Daily Allowance.

KOOGLE PEANUT BUTTER SURPRISE
NUTRITION INFORMATION
PER SERVING

CALORIES = 180

PERCENTAGE OF U.S. RECOMMENDED DAILY ALLOWANCES (U.S. RDA)			
PROTEIN	EXCELLENT	CALCIUM	FAIR
NIACIN	GOOD	RIBOFLAVIN	FAIR
THIAMINE	GOOD	VITAMIN A	NONE
IRON	GOOD	VITAMIN C	NONE

1b. Sample disclosure format using adjectival descriptions of nutrient content.

The Experimental Procedure

Subjects were selected on a quota basis from a group of California residents and tourists participating in a commercial television pilot preview test. During the preview test data were gathered on demographic and socioeconomic variables as well as frequency of purchase, importance of, and last brand purchased of peanut butter. Twelve subjects at a time were asked to stay after the preview to participate in a focus group interview. During this interview they were shown three 30-second commercials -- a control commercial, then the two test commercials (the test commercials were reversed in presentation order during 1/2 of the sessions to guard against any order bias). After exposure to the commercials, subjects responded to a short questionnaire regarding the test commercials including questions about their brand choices and brand preference, their images of the test brands, and their feelings about the decision situation. These responses were analyzed by comparing groups exposed to the information in the two different formats. Where appropriate, the responses were analyzed separately for Skippy (expectancy-discrepant condition) and for Koogle (non-discrepant condition).

Following the experiment, subjects were debriefed. Subjects were given a letter explaining the use of false brand information and giving subjects the real brand data.

The findings reported below and their implications for public policy are discussed in the context of a situation where consumers currently have inaccurate information which needs to be corrected before they can make accurate brand choice decisions. The major concern is with finding an effective method for communicating expectancy-discrepant information in such a manner that consumers come away with accurate brand perceptions. Secondly, interest is in discovering how consumers deal with this discrepant information, i.e. how they feel about their decision situation and their information environment.

Findings and Public Policy Implications
From the Study

Assuming one major role of public policy in protecting consumers is to point out areas where consumers' product knowledge is either insufficient or inaccurate, data from this study indicate adjectival descriptions communicate more effectively with consumers than do percentage contents data.

The More Nutritious Brand

After exposure to nutrition information in one of the two test formats (percentage RDA or adjectival descriptions), subjects were asked to identify the "more nutritious" brand in order to judge the communications effectiveness of the disclosures. An overall chi-square test of the difference between the observed and expected experimental cell frequencies for those subjects choosing each brand as the "more nutritious" one revealed a significant difference between groups exposed to the two different information formats regardless of the amount of information they were exposed to (Tables 2 and 3). Examination of the data suggests that the adjective format is superior to the percentage format in helping respondents identify the "more nutritious" brand.

TABLE 2
Percentage of Respondents Choosing Each Brand as "More Nutritious"

Amount of Information	Choice of "More Nutritious" Brand by Information Format					
	Percentage		Adjective			
	(n)	Koogle	Skippy	(n)	Koogle	Skippy
Control Group No Nutrient Information	28 ^a	17.8	82.2	30	0.0	100.0
Four Nutrients plus Calories	60	10.0	90.0	60	32.2	67.8
Eight Nutrients plus Calories	60	11.7	88.3	59 ^a	30.5	69.5

^aThe 60 respondents in the original control group were randomly assigned to one of two 30-respondent control groups each of which was used as a control for one of the two information format manipulations. The sample for the percentage format control group contained 28 respondents and the sample exposed to information on eight nutrients via the adjective format contained 59 respondents in this case (three of the respondents gave no response to the question under analysis).

TABLE 3
Chi-Square Analysis of Respondents Choosing Each Brand as "More Nutritious" after Receiving Nutrition Information in Various Formats

Brand Chosen as "More Nutritious"	Information Format	
	Percentage	Adjective
Koogle	18	37
Skippy	130	112
Totals	148 ^a	149 ^a

$\chi^2 = 7.08$; d.f. = 1; $p < .01$.

^aTwo respondents in the percentage group and one respondent in the adjective group gave no answer to this question.

Brand Image In Relation to an Ideal Brand

After exposure to the nutrition information, subjects were queried regarding the perceived distance of each of the test brands from their "ideal" brand. Subjects were simply asked to indicate how close each of the test brands was to their "ideal" brand on an overall basis rather than attribute by attribute, the method that has been used in several past studies (Jacoby, Speller, and

Kohn, 1974a, b). Theoretically, if nutrition is an important attribute and a brand is perceived as nutritious, it should be rated as "close to" the ideal. Conversely, the "less nutritious" brand should be seen as "far from" the ideal. Thus, it was expected that, following exposure to the most effective format for the nutrition information, Koogle would be seen as closer to the ideal and Skippy as further from the ideal. It appeared that for Koogle, the unknown brand for which the information presented to subjects was not discrepant with pre-exposure knowledge, the nutrition information neither moved the brand closer to nor further from the ideal brand no matter what the format of the information. Analysis of variance of the mean perceived distance scores between groups revealed a non-significant main effect for format (Tables 4 and 5).

TABLE 4
Perceived Distance of Koogle from Ideal Brand^a

Information Format:	(n)	Per-centage	(n)	Adjec-tive	Marginal Means
<u>Amount of Information:</u>					
Control Group -- No Nutrient Information	20	3.12	20	3.65	3.39
Four Nutrients plus Calories	60	3.68	60	3.22	3.45
Eight Nutrients plus Calories	60	3.20	60	3.54	3.36
<u>Marginal Means</u>	<u>140</u>	<u>3.49</u>	<u>140</u>	<u>3.42</u>	<u>3.42</u>

^aPerceived distance from ideal was measured on a 5-point scale from 1 = just like my ideal to 5 = extremely far from my ideal.

TABLE 5
Analysis of Variance of Perceived Distance
of Koogle from Ideal Brand

Source of Variation	d.f.	Mean Square	F Ratio
<u>Main Effects</u>			
Amount of Information	2	.24	.17
Format of Information	1	.05	.04
<u>Interaction Effects</u>			
Amount X Format	2	6.59	4.68**
Residual (Error Term)	274	2.00	
<u>Total</u>	<u>279</u>		

**p = .01

However, the information presented to respondents on Skippy, the well-known brand, disconfirmed subjects' preconceived notions and/or expectations about the brand. Respondents entered the experiment with a strong pre-exposure preference for Skippy.³ The disclosure information portrayed Skippy as a "less valuable" product than subjects had expected. Analysis of variance of the mean perceived distance score of Skippy from the "ideal" brand revealed a significant difference between groups exposed to the two different information formats regardless of the amount of information they were exposed to. Skippy was perceived as further from the "ideal" by those exposed to adjectival information than by those who saw percentage information (Tables 6 and 7).

TABLE 6
Perceived Distance of Skippy from Ideal Brand^a

Information Format:	(n)	Per-centage	(n)	Adjec-tive	Marginal Means
<u>Amount of Information:</u>					
Control Group -- No Nutrient Information	26	1.85	28	1.70	1.77
Four Nutrients plus Calories	60	1.77	60	2.20	1.98
Eight Nutrients plus Calories	60	1.82	60	2.47	2.14
<u>Marginal Means</u>	<u>146</u>	<u>1.81</u>	<u>148</u>	<u>2.15</u>	<u>2.32</u>

^aPerceived distance from ideal brand was measured on a 5-point scale from 1 = just like my ideal to 5 = extremely far from my ideal.

TABLE 7
Analysis of Variance of Perceived Distance
of Skippy from Ideal Brand

Source of Variation:	d.f.	Mean Square	F Ratio
<u>Main Effects</u>			
Amount of Information	2	2.86	2.99*
Format of Information	1	12.34	12.90***
<u>Interaction Effect</u>			
Amount X Format	2	3.00	3.14*
Residual (Error Term)	288	.96	
<u>Total</u>	<u>293</u>		

*** p = .001, * p = .05.

These data suggest that the FTC and other regulatory agencies may want to develop different information disclosure requirements for new and for established products. However, in recommending the use of adjectival disclosures, the regulatory agencies must be cautious that their proscriptions and remedies are not punitive in nature. The danger with using adjectival disclosures is that they appear to facilitate downward evaluations of products.

Consumers' Psychological Feelings

A related problem arises when considering the effects of expectancy-discrepant information (i.e., disclosures designed to correct an erroneous perception of a well-known brand) on consumers' subjective-psychological feelings. How can consumers be expected to deal with the resulting conflict, discomfort and dissatisfaction? Will they abandon any thoughts of purchase of the brand described? Will they continue buying the brand they have always bought? Will they stop buying that type of product altogether assuming one brand is just like the others? Or will they seek additional information to quiet their discomfort and dissatisfaction?

³Using the responses of subjects in the control group that saw no commercial messages as surrogate indicators for the experimental groups' pre-exposure brand evaluations showed 93.7% of the respondents preferred Skippy over Koogle.

Subjects in this test were asked to indicate whether they preferred Skippy or Koogle. They then were asked to rate their certainty regarding their choice, their confusion during the decision-making process, their satisfaction with their decision, and their desire for more information. These ratings were on 5-point scales with 1 representing the most positive response.

By comparing the experimental groups on their mean scores on these subjective state scales, analysis of variance revealed significant differences for satisfaction and desire for more information (Tables 8 and 9) between the groups exposed to the two different formats.

TABLE 8
Mean Subjective State Scores

Amount of Information: Variable	Information Format					
	Percentage			Adjective		
	4	8		4	8	
	Control	Nutri-ents	Nutri-ents	Control	Nutri-ents	Nutri-ents
Satisfaction ^a	1.30	1.43	1.58	1.20	1.72	1.75
Certainty ^b	1.33	1.53	1.63	1.27	1.82	1.63
Confusion ^c	1.30	1.38	1.50	1.30	1.65	1.53
More information ^d	2.33	2.50	2.32	2.37	2.87	2.82

^aFive-point scale: 1 = very satisfied; 5 = very dissatisfied.

^bFive-point scale: 1 = very certain; 5 = very uncertain.

^cFive-point scale: 1 = not confused; 5 = very confused.

^dFive-point scale: 1 = too much information already; 5 = great deal more.

TABLE 9
Summary of the Significant Effects of Information on Respondents' Subjective States as Revealed by Analysis of Variance

Variable	Analysis of Variance					
	Main Effects				Interaction	
	Amount		Format		Amount X Format	
	CG ^a	NCG ^b	CG	NCG	CG	NCG
Satisfaction	**	--	*	*	--	--
Certainty	**	--	--	--	--	--
Confusion	--	--	--	--	--	--
Desire for More Information	--	--	**	**	--	--

^aAnalysis of variance performed including the control group.

^bAnalysis of variance performed excluding the control group.

-- = not significant ** p = .01 * p = .05

The data from this study suggest that those respondents receiving adjectival disclosures are less satisfied with their choices than are those receiving percentage disclosures. This decreased satisfaction appears to result in a desire for more additional information, probably to clarify the discrepancy evident between the new information and prior brand expectations. If this desire for additional information leads to actual search for information, it may not be too unreasonable to expect consumers to seek additional information. Consumers may look

for information to "validate" the adjectival labels attached to various products. Once validated, these labels might be used with more confidence in the future.

Summary

In designing legislation and judicial remedies to protect consumers, major concern of policy makers lies in the prevention of deception or the correction of images based on deceptive or misleading information. Both of these objectives can be obtained through communication of the true facts about brands on the market. Policy may thus require the communication of information discrepant with commonly held beliefs about a brand.

This study attempted to determine which of two alternative information disclosure formats would be most effective in communicating expectancy-discrepant information to consumers. The data indicate that an adjectival format is superior to a percentage format in helping respondents identify the "more nutritious" brand, in increasing the perceived distance between the "less nutritious" brand and the "ideal" brand, and in creating a lower level of satisfaction with the chosen brand which appears to be related to a desire for more information. It appears that with respect to the effects hypothesized in this study, exposure to expectancy-discrepant information may cause consumers to reevaluate their attitudes toward a brand or, at least, desire further information to clarify the discrepancy revealed. Both of these effects are desirable from a consumer protection point of view when considering a situation where currently held consumer beliefs are based on erroneous information.

Further research is needed to evaluate the effects of expectancy-discrepant information on consumers and to evaluate the impact of adjectival descriptors over time.

References

- Rolph E. Anderson, "Consumer Dissatisfaction: The Effect of Disconfirmed Expectancy on Perceived Product Performance," *Journal of Marketing Research*, 10 (February, 1973), 38-44.
- James R. Bettman, "Issues in Designing Consumer Information Environments," *Journal of Consumer Research*, 2 (December, 1975), 169-177.
- Richard N. Cardozo, "An Experimental Study of Customer Effort, Expectation and Satisfaction," *Journal of Marketing Research*, 2 (August, 1965), 244-249.
- Jacob Jacoby, Donald E. Speller, and Carol A. Kohn, "Brand Choice Behavior as a Function of Information Load," *Journal of Marketing Research*, 11 (February, 1974), 63-69.
- Jacob Jacoby, Donald E. Speller, and Carol A. Kohn, "Brand Choice Behavior as a Function of Information Load: Replication and Extension," *Journal of Consumer Research*, 1 (June, 1974b), 33-42.
- J.C. Maloney, "Curiosity Versus Disbelief in Advertising," *Journal of Advertising Research*, 2 (June, 1962), 2-8.
- Nutrition Advertising Rule*. 39 Fed. Reg. 39482, 1974.
- Richard W. Olshavsky and John A. Miller, "Consumer Expectations, Product Performance and Perceived Product Quality," *Journal of Marketing Research*, 9 (February, 1972), 19-21.

J.E. Russo, "More Information is Better: A Reevaluation of Jacoby, Speller and Kohn," Journal of Consumer Research, 1 (January, 1975), 68-72.

J.E. Russo, G. Krieser, and S. Miyashita, "An Effective Display of Unit Price Information," Journal of Marketing, 39 (April, 1975), 11-19.

Debra L. Scammon, "The Effects on Consumers of Varying the Amount and Format of Purchase Relevant Information," unpublished doctoral dissertation, University of California, Los Angeles, 1976.

SHOPPING TIME AND LEISURE TIME: SOME PRELIMINARY
CROSS-CULTURAL COMPARISONS OF TIME-BUDGET EXPENDITURES

Douglass K. Hawes, University of Wyoming
Sigmund Grønmo, Norwegian Fund for Market and Distribution Research
Johan Arndt, Norwegian School of Economics and Business Administration

Abstract

The findings reported herein represent an attempt to consumers' use of time for shopping, and the impact of shopping time on other categories of time expenditure. The results showed great similarity in the time spent on shopping in Norway and in the United States, but wide differences in the use of time for other activities.

Introduction

Two papers presented at the 1976 Association for Consumer Research Annual Conference dealt with the topic of time budgets in consumer behavior (Arndt and Grønmo, 1977; Hawes, 1977). As pointed out in both of these papers, the subject of time budgets and their impact on consumption behavior has been an underresearched area (Jacoby, et al., 1976; Arndt, 1976). It is undeniable that time underlies all human activities--in one way or another, the clock dominates behavior. While a person may add to his money budget and, at least in the short run, his energy budget, one's time budget is immutably fixed at 168 hours per week.

This brief research report builds upon the two aforementioned 1976 ACR papers and presents a comparison of selected time budget expenditures in two countries--Norway and the United States. One does not have to read very deeply into the consumer/buyer behavior literature to realize the paucity of cross-cultural comparative consumer behavior studies.

A growing interest in this type of study can, however, be identified (Plummer, 1977). Without the knowledge of consumer behavior in different cultures it is difficult to develop general theories of this behavior. The lack of knowledge in this area has led to inefficiencies in resource allocations as well as socio-political conflicts. These effects have been strengthened by the growth in the operations of multinational corporations in various parts of the world (Sheth & Sethi, 1973). Neither the concept of the "universality of cultural traits" nor the concept of the "distinctiveness of each culture" seems to be a satisfactory approach in the study of consumer behavior. In this area, the question rather is: what is universal and what is distinct? It is on this basis that this paper hopes to make a contribution.

Background

The reader is referred to the papers by Arndt and Grønmo (1977) and Hawes (1977) for a detailed discussion of the earlier research streams which led to those two papers. Briefly, Arndt and Grønmo posit a "time expenditure approach to consumer behavior...viewing consumer behavior as a household management area in which members of the household interact and allocate tasks to maximize satisfaction of the input of the scarce resource time (or its equivalent money)" (p. 230). They develop a theoretical paradigm which positions the variable "Orientation to Shopping" "as a hypothetical construct intervening between 'Time Spent on Shopping' and a set of antecedent variables believed to be correlates of 'Orientation to

Shopping'" (p. 230). The authors go on to suggest that certain activities are less amenable to contraction or expansion in time consumption than other activities, given an autonomous change in another activity, such as shopping. The authors then investigate the accuracy of the model using Norwegian data, including time expenditure elasticity for those with "long" shopping time versus those with "short" shopping time.

Hawes focused on leisure-time (truly, discretionary time) expenditures and presented American data on nine categories of leisure-time pursuits cross-classified by demographics. His approach was more descriptive, atheoretical, and applied, and approached the usage of time from a demographics-based market segmentation perspective. He then presented some data on male and female respondents' inclinations about what they would do with a) an extra two hours in the day, and b) with a three-day weekend every week.

While leisure-time patterns have been used as a basis for measuring life style, shopping-time has been considered as an important dimension of shopping behavior. For several reasons we may expect that a relationship between these two variables exists:

- Shopping may influence the duration of leisure-time. For instance, purchase of goods may be a means to reduce the time spent on housework. Besides, the amount of time spent on shopping affects the amount of time available for leisure. Conversely, the amount of time available for shopping may be affected by variations in leisure-time.
- Shopping may influence the "content" of leisure-time. How the leisure-time is used, depends to some degree on consumption and purchase of goods.
- On the other hand, shopping may be influenced by the "content" of leisure-time. For instance, TV-watching and newspaper reading may stimulate shopping because of the advertisements in the media.
- Some people find a leisure component in shopping, which may be perceived as recreation, entertainment or a socializing activity.

Focus of This Paper

Given a) the findings of these two aforementioned papers, b) the potential of cross-cultural comparative research in aiding in the development of comprehensive explanatory models of consumer choice behavior (Arndt, 1976), and c) the approximate, but feasible, comparability of the two data bases involved--it was decided to examine the differences in various categories of time expenditures for Americans spending "long" or "short" amounts of time shopping, and the expenditures of time by Norwegians for various leisure-time pursuits. Therefore, the following two research questions are addressed in this paper:

1. How does the time spent by Norwegians on shopping and on selected leisure-time pursuits compare with the time spent by Americans on the same categories of activities when both sets of respondents are segregated by the same or similar demographic categories?
2. How does the time spent by Americans on selected leisure-time activities vary with time spent shopping in comparison with Norwegians? "Short" shopping time is defined as less than or equal to 30 minutes per day (3.5 hours per week); "long" shopping time is defined as more than 30 minutes per day for both data bases.

In both life style and shopping behavior we find variations between groups and nations. One purpose of this analysis, then, is to study the relationship between leisure-time patterns and shopping-time in various social (demographic) groups in the USA and in Norway.

Data Bases

The source of the Norwegian data is the Time Budget Survey conducted by the Norwegian Central Bureau of Statistics in 1971-72. The design of the survey and other findings are presented in Statistisk Sentralbyrå (1974, 1975b) and Grønmo (1976). A total of 3,040 persons completed a time usage diary and participated in before-and-after interviews. This figure represents a response rate of 58 percent. Only the married persons in the Norwegian sample are included in the following analysis. This is done in order to maximize the comparability between the Norwegian and the American data. Further details on the methodology are outlined in Arndt and Grønmo (1977).

The source of the American data was a nationwide sample of 1000 households participating in Market Facts Inc. (Chicago) Consumer Mail Panel. Two thousand identical, 20-page, color-coded questionnaires were sent out to these households in the summer of 1973, so that both the male and female head-of-household could independently respond. There were a total of 1115 usable responses, or a response rate of 56 percent. Further details on this methodology are presented in Hawes (1977).

Both surveys collected demographic data including respondent's age, education, total annual household income, occupation (males) or employment status (females), household size, and population density/degree of urbanization. The number of hours the respondent spent at work each week was also determined.

Results

Based upon the respondents in the Norwegian and the American surveys, some interesting similarities and differences appear, reflecting differences in overall life-style between the two countries. Investigation is continuing into the social and cultural bases for these similarities and differences. The following discussion will focus on some of the more interesting similarities and differences.

Female Respondents

Tables 1 and 2 contrast the time spent on shopping and on various leisure-time pursuits by Norwegian and American women. The first similarity of note is the approximately equal amount of time spent shopping by females in both countries. The same basic patterns hold across the demographic variable categories in both countries with, for example, the greatest amount of time in the respective category being spent by women aged 35-44, those who

are full-time homemakers, and those who are better educated. There is a statistically significant drop in hours spent shopping by Norwegian women in rural areas (versus central city areas), and among those who work 45 hours per week or more (versus 29 hours per week or less), that does not show up in the American data.¹ This is probably a function of differences in geography, transportation, retailing dispersion and work/living locations between the two countries.

Overall, Norwegian women spend from one-half to about three-quarters as much time playing with their children as do American women. The differences are very great when examined in relation to the hours per week spent at work. In the extreme categories of this variable, American women devote from four to almost thirteen times as much time to their children than do Norwegian women. Interestingly, while there is a significant drop with increasing income in the number of hours spent with children by American women, just the reverse pattern holds in Norway.

With the exception of those in the lowest income category and retirees, Norwegian women spend about half as much time reading newspapers and magazines as do American women. In these two particular categories, respondents in both surveys spent about the same amount of time on this activity. As one might expect, in both countries the greatest amount of time (in the demographic category) spent reading newspapers and magazines is in the smallest family size (2), and among those working the fewest hours per week in their job.

Norwegian women watch approximately one-third to one-half as much television as American women. Interestingly, while television viewing among the American respondents peaks in the under-25 age category and declines through the 35-44 category, in Norway, viewing increases with age and peaks in the 55-and-over category. These differences may be explained by variations in the number of TV-programs available. While Americans can choose among several programs all day, Norwegians have no such choice between parallel programs. Furthermore, in Norway TV-programs are transmitted during only part of the day. In both countries, viewing is least (across education levels) among those with a college education, declines with increasing family size, and shows no significant difference between the two extreme categories of hours per week spent at work.

Norwegian women spend approximately one-fifth to one-half as many hours per week on hobbies, games and crafts as do American women. There is a much greater decline in these activities with increasing income in Norway than in the United States. In both samples, the greatest amount of time is spent by those in the oldest age category, and increasing education is associated with a decrease in the amount of time spent on these activities. Interestingly, participation in these activities reaches a minima around 35-40 hours per week of employment and increases on both sides of this range. There appears to be a phenomenon, more apparent in the United States data, that women who work long hours also find time for a number of other activities. Whether these women are merely "hyper," have great energy and involvement, or are basically "workaholics" cannot be determined from the small sample involved.

Norwegian women spend a great deal more time (two to

¹"Significant" refers to differences significant at the .05 level or greater (two-tailed) in the Norwegian data, and .1 level or greater in the American data. The differences in sample size suggest the differences in cut-off level of significance. In many cases, the level of significance far exceeded the .05/.1 level.

TABLE 1

Hours Per Week Spent in Selected Time Budget Categories by Demographics--
Females--Norwegian Sample

	N	Shopping	Playing with or helping your children	Reading newspapers and magazines	Watching television	Hobbies, games, crafts, etc.	Visiting with friends or relatives	Participating in sports or athletics	Attending sporting events as a spectator	Entertainment outside the home (other than sporting events)
Age										
Less than 25	(125) ^b	3.4 ^a	11.4	1.8	5.3	1.2	18.1	3.7	.1	1.1
25 - 34	(291)	4.1	10.1	2.2	5.8	.8	15.0	3.4	.1	.8
35 - 44	(237)	4.2	5.5	2.3	6.6	1.0	13.5	2.8	0	.2
45 - 54	(247)	3.8	1.2	2.6	6.7	1.4	13.2	2.9	0	.5
55 and over	(264)	3.1	.6	3.0	7.1	3.3	12.7	2.5	0	.1
Education										
Public school (7 years)	(382)	3.3	2.5	2.5	6.8	2.1	12.9	2.4	0	.2
Public school (9-10 years)	(632)	3.8	6.3	2.3	6.3	1.3	14.6	3.2	0	.6
High school	(78)	4.3	6.8	3.2	6.5	1.1	16.6	2.5	0	.6
College	(71)	4.3	9.0	2.8	5.2	1.2	14.2	4.6	0	.8
Total Household Income										
Less than Nkr. 15,000	(21)	3.1	3.2	4.0	8.2	4.0	11.5	3.0	0	.1
Nkr. 15,000 - 29,000	(229)	2.6	3.8	2.7	6.7	2.4	13.3	1.8	0	.1
Nkr. 30,000 - 44,000	(410)	3.7	6.2	2.3	6.5	1.4	14.6	3.0	.1	.4
Nkr. 45,000 or more	(394)	4.5	5.5	2.4	6.2	1.0	14.8	3.7	0	.9
Employment Status										
Works full-time	(211)	2.6	2.6	1.9	6.3	1.1	11.9	3.0	0	.6
Works part-time	(255)	3.7	3.0	2.4	6.3	1.4	13.4	3.1	0	.3
Student	(9)	6.4	7.0	1.3	2.0	.7	17.7	8.6	0	.9
Pensioner	(36)	2.5	.6	4.0	7.5	4.0	12.9	1.8	0	.3
Full-time homemaker	(651)	4.2	7.2	2.6	6.5	1.7	15.2	2.9	0	.5
Household Size										
2	(319)	3.2	.5	2.9	6.7	2.6	14.2	3.1	0	.4
3	(252)	3.8	4.4	2.5	6.2	1.8	14.7	3.2	0	.5
4	(299)	4.1	7.4	2.2	6.6	.9	13.8	2.7	0	.6
5 or more	(283)	3.9	8.8	2.2	6.0	1.0	13.8	2.9	.1	.4
Population Density/Degree of Urbanization										
Rural	(528)	3.0	5.1	2.3	6.0	1.7	13.9	2.5	.1	.3
Urban	(432)	4.0	5.8	2.6	6.6	1.4	14.5	3.2	0	.3
Central City	(205)	5.2	4.5	2.4	6.9	1.4	14.0	3.8	0	1.2
Hours Per Week Spent at Work										
29 hrs. or less	(234)	3.8	3.2	2.6	6.2	1.4	13.9	3.2	0	.3
30 - 39 hrs.	(83)	3.2	2.6	1.8	6.9	1.2	13.1	3.9	0	.6
40 - 44 hrs.	(119)	2.3	2.8	1.6	5.9	.8	10.7	2.3	0	.5
45 hrs. or more	(27)	1.7	.5	1.6	6.8	1.2	11.5	2.1	.3	.2

^aCell values are mean hours-per-week "on the average." The standard deviation of each mean is not listed in the interest of minimizing obfuscation.

^bNumber of respondents in the demographic category segment.

three times as much) visiting friends or relatives than do American women. In both countries, visiting is highest in the youngest age category and least among college graduates, and is little affected by the hours per week spent at work. In the United States sample, visiting dropped significantly between families with one child and those with three or more; in Norway, family size had no significant affect.

It is in the area of sports and athletics that the greatest differences between the two countries emerge: very simply, Norwegians are "doers" and Americans are spectators. Norwegian women spend from two to four times as much time participating in sports or athletics than American women. While the American women spend only .6 to 1.1 hours per week as a spectator, this figure represents three to five times as much time as the expenditures of Norwegian women. While the data from both countries shows the same basic age-participation pattern, there is a much greater proportional increase in participation among Norwegian women with a college degree over those with only high school degrees than is the case in the

United States. In both countries, participation is not significantly affected by family size or total household income.

Women in the United States sample indicated that they spend much more time in non-sports-related entertainment outside the home--from three to more than seven times as much--than their Norwegian counterparts. In both countries, it is the youngest, best educated, highest income women who go out the most in comparison with those in the other categories of these demographic variables. Larger families do not deter women from going out in either country. Again, American women who work long hours are much more likely than their Norwegian counterparts to find time for this type of leisure-time activity.

Male Respondents

By and large, it is evident from Tables 3 and 4 that the same patterns hold in comparing the time expenditures of Norwegian and American males as was found in comparing

TABLE 2

Hours Per Week Spent in Selected Time Budget Categories by Demographics--
Females--United States Sample

	N	Shopping	Playing with or helping your children	Reading newspapers and magazines	Watching television	Hobbies, games, crafts etc.	Visiting with friends or relatives	Participating in sports or athletics	Attending sporting events as a spectator	Entertainment outside the home (Other than sporting events)
Age										
Less than 25	(127) ^b	4.3	18.4 ^a	3.7	18.6	5.1	9.0	1.7	1.1	3.9
25 - 34	(140)	3.7	14.9	4.1	16.2	5.1	5.6	1.1	.7	3.2
34 - 44	(92)	4.7	9.4	4.9	12.5	5.4	5.1	1.7	.9	3.4
45 - 54	(104)	4.2	5.1	5.7	13.8	5.4	4.5	1.4	.6	3.2
55 and over	(140)	4.2	2.1	7.3	15.9	6.9	5.1	.6	.4	2.5
Education										
Less than high school graduate	(122)	4.3	8.1	4.6	18.3	5.6	5.8	1.4	1.1	3.3
High school graduate	(268)	4.0	11.1	5.2	16.5	5.5	6.2	1.3	.6	2.9
Some college	(145)	4.4	10.8	5.3	13.1	5.9	5.9	1.0	.6	3.5
College graduate or post-graduate	(65)	4.1	9.1	5.7	12.6	5.2	5.4	1.2	.7	3.8
Total Household Income										
\$4,000. - 9,999.	(252)	4.3	11.7	4.7	18.0	6.0	6.8	1.5	.8	3.1
\$10,000. - 14,999.	(199)	4.0	10.7	5.2	15.1	5.1	5.3	.9	.6	2.8
\$15,000. or more	(152)	4.2	7.1	5.8	12.5	5.8	5.5	1.4	.6	4.0
Employment Status										
Works full-time	(156)	3.6	4.5	4.2	12.5	4.2	4.4	1.1	.9	3.2
Works part-time; is self-employed is not employed because is retired, student, disabled, etc.	(117)	3.8	9.0	5.7	14.3	6.3	6.5	1.7	.9	4.4
Full-time homemaker	(319)	4.5	13.5	5.4	17.4	6.0	6.6	1.2	.6	2.9
Household Size										
2	(250)	4.1	3.2	5.8	16.2	5.9	5.9	1.0	.5	3.3
3	(143)	4.5	16.4	4.3	16.6	5.9	7.1	1.7	1.1	3.4
4	(162)	3.8	15.8	4.9	15.1	5.1	5.5	1.3	.7	3.1
5 or more	(48)	5.0	8.9	4.7	11.8	5.0	4.1	.9	.7	2.9
Population Density/Degree of Urbanisation										
Rural	(147)	4.2	10.2	4.5	14.5	4.8	5.8	1.1	.8	3.0
Urban	(266)	4.4	10.6	5.4	16.8	6.3	6.4	1.4	.7	3.4
Central City	(190)	3.8	9.7	5.2	15.0	5.3	5.4	1.1	.7	3.1
Hours Per Week Spent at Work^c										
24 hrs. or less	(75)	3.5	11.3	5.2	13.8	4.6	5.5	1.7	.6	3.0
25 - 34 hrs.	(36)	3.9	7.7	4.7	9.5	5.3	5.5	.7	.5	4.5
35 - 39 hrs.	(42)	3.3	3.8	4.0	11.7	2.8	4.2	.9	.5	4.1
40 - 45 hrs.	(97)	3.4	5.1	3.8	11.6	3.7	4.2	.5	.7	3.1
45 hrs. or more	(31)	3.2	6.4	4.4	12.3	4.2	6.8	2.5	2.1	4.8

^aCell values are mean hours-per-week "on the average." The standard deviation of each mean is not listed in the interest of minimizing obfuscation.

^bNumber of respondents in the demographic category segment.

^cHours per week refers to hours spent at "primary job." Only those respondents with one job are included here.

females from the two countries. Although Norwegian men spend slightly more time per week shopping than their American counterparts, the absolute values are very similar. There is some indication that college educated Norwegian men spend disproportionately more time shopping than American college educated men. Also, time spent shopping holds up much better in Norway among those working 50 or more hours per week, than it does in the United States.

American men spend from approximately two to five times as much time playing with their children as do Norwegian men. In examining education and income, one notes that in Norway, the greatest amount of time with children is spent by those in the highest income and education categories; in America, the most time is spent by those in low-to-moderate income groups and with a high school/some college education. In both countries, the amount of time spent playing with children does not increase signifi-

cantly as more than one child enters the family. As one might expect, rural males in both countries spend the least amount of time playing with their children--probably because there is a less clear-cut distinction between working with and playing with the children.

Norwegian men, like their wives, spend about two-thirds as much time reading newspapers and magazines as their American counterparts. This holds except in the lowest income categories and among retirees, where the time spent reading is quite similar in the two countries. Across income categories, the greatest amount of time in this activity in the United States is in the highest income bracket; in Norway, it is in the lowest income bracket. There are wide differences across specific occupational categories in both countries.

Norwegian men are again similar to their wives, in that they watch about one-half to two-thirds as much televi-

TABLE 3

Hours Per Week Spent in Selected Time Budget Categories by Demographics--
Males--Norwegian Sample

	N	Shopping	Playing with or helping your children	Reading newspapers and magazines	Watching television	Hobbies, games, crafts, etc.	Visiting with friends or relatives	Participating in sports or athletics	Attending sporting events as a spectator	Entertainment outside the home (Other than sporting events)
Age										
Less than 25	(69) ^b	1.6 ^a	2.5	2.6	7.1	1.4	15.8	3.4	.2	.8
25 - 34	(202)	2.5	3.5	2.4	8.1	.7	13.9	3.3	.2	.8
35 - 44	(238)	2.4	2.2	2.7	8.6	1.2	11.5	4.2	.1	.6
45 - 54	(252)	2.3	.5	3.2	8.5	1.0	10.2	3.7	.2	.3
55 and over	(293)	2.7	.3	4.4	9.4	1.1	9.7	3.6	0	.4
Education										
Public school (7 years)	(326)	2.6	.7	3.6	9.5	1.0	11.1	3.0	.1	.4
Public school (9-10 years)	(544)	2.3	1.7	3.0	8.4	1.0	11.5	3.8	.2	.5
High school	(64)	1.8	1.7	3.3	8.8	1.4	12.7	5.9	.3	.8
College	(122)	2.7	2.8	3.3	6.9	1.0	11.7	3.7	.1	.5
Total Household Income										
Less than Nkr. 15,000	(36)	2.0	.5	4.5	8.9	.9	12.2	2.4	0	0
Nkr. 15,000 - 29,000	(398)	2.5	1.1	3.9	8.2	1.1	11.7	3.3	0	.3
Nkr. 30,000 - 44,000	(401)	2.3	1.7	3.1	8.9	1.1	11.1	3.5	.1	.6
Nkr. 45,000 or more	(370)	2.5	1.7	3.0	8.4	1.0	11.7	4.4	.2	.7
Occupation										
Professional/Technical/Managerial/Adm.	(144)	2.6	2.8	3.2	7.0	.9	11.7	3.5	.1	.6
Clerical or Sales	(123)	1.4	1.2	3.0	8.5	1.2	10.6	4.4	.2	.8
Agriculture, Forestry, Fishing	(107)	2.3	.8	3.5	6.5	.4	10.7	1.5	.1	.4
Transport/Communication	(103)	3.0	2.1	2.7	8.1	.9	11.4	3.3	.2	.7
Industry/Construction	(355)	2.1	1.4	3.0	9.2	1.0	11.5	4.2	.2	.4
Service/Military	(37)	2.0	.8	2.7	7.4	.7	12.4	3.8	.2	1.0
Student	(8)	.6	2.4	4.5	7.5	1.3	21.1	2.0	.1	.1
Pensioner/Non-Employed	(182)	3.5	1.0	4.3	10.5	1.6	11.6	3.9	0	.4
Household Size										
2	(281)	2.6	.2	4.1	9.4	1.6	10.5	3.3	.1	.4
3	(227)	2.1	1.3	3.9	8.0	.7	12.3	4.0	.1	.6
4	(283)	2.5	2.3	2.6	8.2	.8	12.2	4.4	.2	.6
5 or more	(260)	2.3	2.3	2.4	8.6	.9	10.9	3.0	.1	.5
Population Density/Degree of Urbanization										
Rural	(436)	2.4	1.2	3.4	8.2	1.0	11.7	3.1	.1	.4
Urban	(426)	2.5	1.6	3.1	8.9	1.1	11.4	3.9	.2	.5
Central City	(196)	2.3	2.0	3.3	8.8	1.0	11.1	4.5	.2	.7
Hours Per Week Spent at Work										
39 hrs. or less	(117)	2.3	1.9	3.5	7.6	.8	12.6	3.7	0	.4
40 - 44 hrs.	(532)	2.2	1.6	3.0	8.8	1.1	11.2	4.2	.1	.6
45 - 49 hrs.	(99)	2.1	1.6	3.2	7.7	.5	10.2	2.4	.2	.3
50 hrs. or more	(172)	2.5	1.7	2.8	7.0	.8	11.0	2.9	.1	.7

^aCell values are mean hours-per-week "on the average." The standard deviation of each mean is not listed in the interest of minimizing obfuscation.

^bNumber of respondents in the demographic category segment.

sion as American men. While in Norway, the least amount of time in front of the TV set is spent by those under 25; in America, both those under 25 and those over 55 spend the most time watching television. In both countries, those with a college education watch the least television, as do those with one child and those who work more than 50 hours per week. There is a significant drop in television viewing with increasing income in the United States; in Norway there is no significant difference across income categories.

Overall, American men spend from three to five times as much time on hobbies, games and crafts than do Norwegian men. There is a significant increase in this activity in the United States as one moves into the 55 and over age category; in Norway there is no significant age effect. In both countries, clerical/sales workers are relatively more heavily into this activity than any other type of

active worker, and there is a decrease in time spent on this activity with increasing family size. The hours per week spent at work also has an effect on time for hobbies, etc. in both countries. Central city residents in the United States spend significantly less time on this activity than do "urban" (suburban) residents; in Norway, there is no effect due to location of residence.

As one might expect after examining the data on Norwegian women, Norwegian men also spend from two to three times as much time visiting friends or relatives than do American men. In both countries, visitation time is greatest among those under 25 and among those in the lowest income categories. In Norway, however, there is a distinct drop among those over 55, while in America there is an increase as one moves into this age category. Hours per week spent at work has no significant effect on time spent on this activity in Norway, while it does

TABLE 4

Hours Per Week Spent in Selected Time Budget Categories by Demographics--
Males--United States Sample

	N	Shopping	Playing with or helping your children	Reading newspapers and magazines	Watching television	Hobbies, games, crafts etc.	Visiting with friends or relatives	Participating in sports or athletics	Attending sporting events as a spectator	Entertainment outside the home (Other than sporting events)
Age										
Less than 25	(64) ^b	1.5	8.5	3.5	15.1	4.0	6.5	3.3	1.1	3.7
25 - 34	(150)	1.3	8.0	3.8	14.1	3.7	4.7	2.3	.8	3.8
35 - 44	(63)	1.5	5.4	4.4	13.0	3.5	4.8	2.8	1.1	3.3
45 - 54	(93)	2.1	2.9	5.1	13.4	3.0	3.7	1.2	.7	2.6
55 and over	(142)	2.6	1.3	7.5	15.2	5.4	4.6	.9	.9	2.8
Education										
Less than high school graduate	(137)	2.0	3.7	4.4	14.7	4.0	5.0	1.2	.8	3.5
High school graduate	(164)	1.8	5.8	4.9	14.5	4.3	4.8	2.3	1.0	3.0
Some college	(109)	1.9	6.0	5.5	15.5	3.8	5.0	1.8	.6	3.2
College graduate or post-graduate	(95)	1.7	4.3	5.8	11.7	4.4	4.0	2.3	1.1	3.1
Total Household Income										
\$4,000. - 9,999.	(219)	2.0	5.9	4.9	15.3	4.5	5.4	2.0	.9	3.1
\$10,000. - 14,999.	(172)	1.9	5.0	4.6	14.3	3.9	4.2	1.7	.8	3.1
\$15,000. or more	(121)	1.6	3.2	6.1	12.3	3.6	4.4	2.1	1.0	3.7
Occupation										
Professional	(58)	1.4	6.1	5.0	12.2	3.6	3.8	2.8	1.3	2.9
Managerial/Administrative	(67)	1.9	4.1	5.6	13.0	3.7	3.8	2.1	.8	3.0
Clerical or Sales	(67)	1.5	3.7	4.9	13.7	5.1	5.1	2.4	1.3	3.1
Craftsman/Kindred	(106)	1.6	5.2	4.5	13.4	3.3	5.3	1.6	.9	3.5
Operative	(74)	1.8	7.0	3.3	13.6	3.4	5.6	1.7	.4	3.7
Laborer, Farmer	(63)	1.6	6.2	4.2	14.3	3.1	3.6	1.8	.5	2.5
Household Size										
2	(212)	2.2	1.4	6.1	14.7	4.6	5.2	1.8	.9	3.4
3	(124)	1.8	7.8	4.6	13.5	3.8	5.2	2.2	.6	2.9
4	(137)	1.2	7.5	4.2	14.1	3.9	3.8	2.0	.9	3.3
5 or more	(39)	2.5	6.7	4.7	14.9	2.9	4.0	1.0	1.3	3.3
Population Density/Degree of Urbanization										
Rural	(125)	1.7	4.4	4.3	12.4	3.8	4.5	1.4	.8	2.8
Urban	(231)	1.9	5.1	5.4	14.5	4.8	5.0	2.0	.8	3.5
Central City	(156)	1.9	5.2	5.3	15.4	3.2	4.6	2.1	1.0	3.1
Hours Per Week Spent at Work^c										
39 hrs. or less	(52)	1.8	4.9	5.5	14.1	4.9	4.8	2.3	.9	3.0
40 - 44 hrs.	(190)	2.0	5.5	4.5	14.4	3.9	4.9	2.2	1.0	3.6
45 - 49 hrs.	(70)	1.8	7.0	4.3	13.3	3.5	4.2	2.1	.8	3.3
50 hrs. or more	(107)	.9	4.4	4.8	11.2	3.0	3.5	1.6	.7	2.4

^aCell values are mean hours-per-week "on the average." The standard deviation of each mean is not listed in the interest of minimizing obfuscation.

^bNumber of respondents in the demographic category segment.

^cHours per week refers to hours spent at "primary job." Only those respondents with one job are included here.

in the United States.

While both Norwegian and American men under 25 spend about the same amount of time participating in sports or athletics, by the higher ages Norwegian men are spending three to four times as much time in these pursuits than American men. Overall, Norwegian men spend about one- and one-half to two times as much time participating than do American men. This is about half the difference between Norwegian and American women implying that, relatively speaking, Norwegian women are much more athletically inclined than American women. In both countries, clerical/sales workers are among the heaviest participants, as are those with one or two children, those in central cities, and those with a work week which is not excessive.

As was noted with the women, American men spend much more time at spectator sports--from four to ten times as much--than do Norwegian men. It may be noted that in both

countries clerical/sales workers are into spectator sports as well as participatory sports relatively more heavily than other occupations.

American men spend from five to seven times as much time in "other" entertainment outside the home than do Norwegian men. This difference is consistent with the differences in time spent on entertainment between American and Norwegian women.

Differences in Life Style and Similarity in Shopping Time

In our comments, males and females have been treated separately. The main reason for this is the traditional spousal role differentiation, which is found in both cultures and most demographic categories. Mainly because of the differences in employment status and time spent at work, the total amount of time spent on all activities in Tables 1-4 is smaller for males than females. In particular, married women spend more time shopping

TABLE 5

Relations Between Mean Time Spent on Shopping and
Mean Time Spent on Other Activities--Norwegian Sample

	MALES		FEMALES	
	Spend Less Than 3.5 hrs/wk Shopping	Spend More Than 3.5 hrs/wk Shopping	Spend Less Than 3.5 hrs/wk Shopping	Spend More Than 3.5 hrs/wk Shopping
	hrs/ wk			
Sleeping, napping	59.4	59.8	60.3	57.9 ^a
Eating meals	9.4	11.4 ^a	10.0	10.1
Personal care	5.2	5.1	6.3	5.8
Working at your job and other work-related activity	41.2	19.7 ^a	13.2	7.1 ^a
Commuting to and from work	3.6	1.5 ^a	1.1	.6 ^a
Housework, necessary home maintenance and lawn care	11.8	17.8 ^a	36.0	38.8 ^a
Playing with or helping your children	1.5	1.8	5.0	5.6
Reading newspapers and magazines	3.2	3.6	2.4	2.7 ^a
Watching television	8.5	9.1	6.4	6.3
Hobbies, games, crafts, etc.	1.0	1.4 ^a	1.6	1.4
Visiting with friends or relatives	10.8	14.4 ^a	14.0	14.4
Participating in sports or athletics	3.9	2.8 ^a	3.5	1.9 ^a
Attending sporting events as a spectator	.2	0	.1	0
Entertainment outside the home (other than sporting events)	.5	.7	.4	.7 ^a
Other major activities (Mean Time Spent Shopping)	7.8 (.4)	6.9 (11.9)	7.0 (.6)	4.2 ^a (10.4)
N=	868	190	794	371

^aDifference between "More" and "Less" significant at the .05 level (two-tailed) or greater.

and playing with their children than do married men.

At the same time, however, our analysis indicates very clearly some cross-cultural differences in use of time, which are reflected for both women and men as well as in most of the other demographic categories. Compared to Norwegians, Americans spend more time playing with their children, reading newspapers and magazines, watching TV, engaging in hobbies, games and crafts, and attending sporting events and entertainment. Americans spend less time than Norwegians visiting with friends or relatives, and participating in sports or athletics.

Despite these variations in leisure-time patterns, indicating different life styles, there is a remarkable similarity between the two countries in shopping time. Our second research question focuses upon the relationship between shopping time and time spent on other activities within each country.

Relations Between Shopping Time and Time Spent on Other Activities

Tables 5 and 6 present the mean times spent on sixteen time budget categories by those who spend less than/more than 30 minutes per day. This time break corresponds to

Arndt and Grønmo's (1977) "short shopping time" and "long shopping time" respectively. The 30 minute figure was based on the basic 15 minute interval into which the day was divided in the Norwegian survey for activity recording purposes. Based on structural conditions in Norway, it was felt that breaking on 15 minutes was too short and breaking on 45 or 60 minutes was too long. A sensitivity test was run on both sets of data, breaking at 15 and 45 minutes per day in the Norwegian sample and 2.5, 3.0, 4.0 and 4.5 hours per week in the American sample. The results did not differ markedly from breaking at 30 minutes per day (3.5 hours per week).

Shopping may be considered more or less a mandatory activity, and given that total available time is fixed, "increases or decreases in shopping time will affect the time available for other activities" (Arndt and Grønmo, 1977). On the other hand, shopping time is not necessarily the causative variable. The amount of time available for shopping can be affected by, for instance, the amount of time spent on work. Certain activities are more elastic than others for most people and hence will show more effect for long and short shopping times than others. Arndt and Grønmo (1977, p. 234) found that overall, long shopping time was negatively related to time spent at income-producing work and positively related to

TABLE 6

Relations Between Mean Time Spent on Shopping and
Mean Time Spent on Other Activities--United States Sample

	MALES		FEMALES	
	Spend Less Than 3.5 hrs/wk Shopping	Spend More Than 3.5 hrs/wk Shopping	Spend Less Than 3.5 hrs/wk Shopping	Spend More Than 3.5 hrs/wk Shopping
	hrs/ wk			
Sleeping, napping	51.0	51.6	51.7	50.9
Eating meals	11.4	12.0	11.0	12.5 ^a
Personal care	6.5	7.3 ^a	7.4	9.0 ^a
Working at your job	43.7	27.9 ^a	18.2	12.4 ^a
Commuting to and from work	5.0	4.5	2.2	1.5 ^a
Other work-related activity	3.3	3.7	2.6	2.9
Housework, necessary home maintenance and lawn care	4.8	10.5 ^a	21.2	24.7 ^a
Playing with or helping your children	5.3	3.4 ^a	9.9	10.5
Reading newspapers and magazines	4.8	6.9 ^a	4.6	5.9 ^a
Watching television	13.8	16.5 ^a	14.8	16.6 ^a
Hobbies, games, crafts, etc.	3.8	5.3 ^a	4.8	6.5 ^a
Visiting with friends or relatives	4.4	6.5 ^a	5.0	7.1 ^a
Participating in sports or athletics	2.0	1.5	1.2	1.3
Attending sporting events as a spectator	.9	.8	.6	.9 ^a
Entertainment outside the home (other than sporting events)	3.2	3.7	3.0	3.5
Other major activities	2.1	2.7	2.4	2.0
(Mean Time Spent Shopping)	(.98)	(6.3)	(2.2)	(6.6)
N=	421	83	330	273

^aDifference between "More" and "Less" significant at the .07 level (two-tailed) or greater.

household work and family care, personal needs, and leisure. Both males and females were aggregated in the analysis, however, and the category "leisure activities" subsumed many different pursuits. It would be both conceptually and, hopefully, operationally useful to know how time spent on various activities varied between sexes in both countries and also on specific leisure-time pursuits.

In examining Tables 5 and 6 one sees that for both males and females in both countries, long shopping time (more than 3.5 hours per week) is significantly positively related to time spent on housework, necessary home maintenance, and lawn care, and negatively related to time spent at income producing work (and, in Norway, other work related activity). In the United States, long shopping time is significantly positively related, for both males and females, to time spent on personal care, reading newspapers and magazines, watching television, hobbies/games/crafts, and visiting with friends or relatives.

In Norway, long shopping time is significantly positively related among the males to time spent eating meals, on hobbies, games, and crafts, and visiting with friends or relatives, and negatively related to time spent commuting to and from work, and time spent participating in sports or athletics. For the Norwegian females, long shopping time is significantly positively related to time spent reading newspapers and magazines and entertainment outside the home (other than sporting events), and negatively related to time spent sleeping or napping,

commuting to and from work, participating in sports or athletics, and "other major activities."

Finally, in the United States, long shopping time is significantly positively related among the females to time spent eating meals and time spent attending sporting events as a spectator, and negatively related to time spent commuting to and from work. For the males, long shopping time is negatively related to time spent playing with the children.

These patterns suggest that in both countries the "heavy" shopper (those who spend a long time shopping) spend less time at work and more time at home in relatively sedentary activities (eating, reading, television viewing, hobbies, etc.), or in visiting friends or relatives either at home or away. In the United States these people also spend more time on personal care than the "light" shoppers. In Norway, this sedentary pattern includes less time participating in sports or athletics and in "other major activities." All in all, the picture is one of a segment of the population for whom the social-psychological aspects of shopping (getting out, interacting) are important, and for whom discretionary time may be both more abundant and more of a burden than a blessing. The question for them is how to spend their time, not how to save it.

Conclusion

The results of this comparative study add further support

to the theoretical paradigm proposed by Arndt and Grønmo (1977), namely that time spent on shopping varies between different consumer categories and directly affects the time spent on other activities. The indications are that there may well be a non-economic "Orientation to Shopping" intervening construct between antecedent or determining conditions and actual shopping behavior measured by time spent on shopping. The American data supports the earlier Norwegian findings quite well.

The importance of these findings lies in several areas. First, it is an addition to the small but growing body of comparative, cross-cultural buyer behavior knowledge. Second, it adds further credibility to the caveat that "what works here, may not work over there."

Third, it provides some reference data for evaluating how the increased wealth of Norway (due to the growth in production in the North Sea oil fields) may affect Norwegian consumption patterns. Finally, there are a number of research opportunities implicit in these results, which, basically, only scratch the surface of the similarities and differences between consumers in both countries. There is a glaring need to investigate the "why" question in all of its social and cultural ramifications.

How do these findings change anything? This is difficult to address in absolute terms. Certainly they suggest the importance of considering a time expenditure approach to buyer behavior decision processes. They also suggest that non-traditional, non-economic variables may play a more important role in buyer behavior than previously recognized. Studies of the role of time in buyer behavior have truly been few (Jacoby, et. al., 1976). It is also, of course, difficult to relate to "proprietary" cross-cultural studies filed away in corporate research departments.

Some of the actions suggested by these results have been alluded to above. Other appropriate actions include a) comparing these behavior patterns in Norway with those in other Scandinavian countries, b) analyzing the activities of retail stores designed to facilitate the latent "social" aspects of shopping, c) expanding the cross-national exchange of data. There is a wealth of state-collected secondary data waiting to be examined from the perspective of someone in a different country.

References

Arndt, Johan. "Reflections on Research on Consumer Behavior," in Beverlee B. Anderson (ed.), Advances in Consumer Research, Vol. 3. Cincinnati, Ohio: Association for Consumer Research, 1976, 213-221.

Arndt, Johan and Sigmund Grønmo. "The Time Dimension of Shopping Behavior: Some Empirical Findings," in William D. Perreault, Jr. (ed.), Advances in Consumer Research, Vol. 4. Atlanta, Georgia: Association for Consumer Research, 1977, 230-235.

Grønmo, Sigmund. Innkjøp Og Tidsbruk: Forbrukerinnsett i Norsk Varefordistribusjon. Oslo: Fondet for Markeds- og Distribusjonsforskning, 1976.

Hawes, Douglass K. "Time Budgets and Consumer Leisure-Time Behavior," in William D. Perreault, Jr. (ed.), Advances in Consumer Research, Vol. 4. Atlanta, Georgia: Association for Consumer Research, 1977, 221-229.

Jacoby, Jacob, G. J. Szybillo, and C. K. Berning. "Time and Consumer Behavior: An Interdisciplinary Overview," The Journal of Consumer Research 2 (March, 1976), 320-339.

Plummer, Joseph T. "Consumer Focus in Cross-National Research," Journal of Advertising 6 (Spring, 1977), 5-15.

Sheth, Jagdish N. and Prakash Sethi. Theory of Cross-Cultural Buyer Behavior. Faculty working paper, College of Commerce and Business Administration, University of Illinois, 1973.

Statistisk Sentralbyrå, Tidsnyttingsundersøkelsen 1971-72. Hefte 2. Oslo: Statistisk Sentralbyrå, 1974; 1975b (Hefte 1).

PERSONAL VALUES: A CROSS CULTURAL ASSESSMENT OF SELF VALUES
AND VALUES ATTRIBUTED TO A DISTANT CULTURAL STEREOTYPE

J. Michael Munson, University of Santa Clara
Shelby H. McIntyre, University of Santa Clara

Abstract

This research investigates the potential of a particular value assessing instrument, the Rokeach Value Survey (RVS), to distinguish cultural differences in personal value structures. Results suggest personal values can be productively employed by international marketers to discriminate peoples of culturally diverse backgrounds. A discriminant analysis disclosed significant differences in the self values held by people from three cultures: Thailand, Mexico and the United States. Additionally, the stereotype held of an average American's personal values differed significantly across the three cultures. The discriminant functions derived from responses to the RVS were able to correctly classify 65% of the people according to their cultural background as compared to a chance level of 22%.

Introduction

The determination of cultural differences is an important element in formulating international marketing strategy. In view of the increasing participation of American corporations in the international market place, and the threat of growing competition from foreign multinational corporations, knowledge and identification of culturally related differences in buyer behavior is a critical concern for marketing research. A fundamental area where cultural differences should be identified is in the culturally generalizable aspects of personality, attitudes, and personal values. Until recently, there has been little empirical research devoted to the analysis of values and their relation to consumer behavior. Research which has been done has focused solely on the values held by people in the United States (Scott and Lamont, 1974; Vinson and Munson, 1976; Henry, 1976; Vinson, Scott and Lamont, 1977).

While it is a widely accepted notion that cross cultural differences in values do exist, there is little empirical evidence actually documenting such differences or suggesting the particular form or pattern they take. In order for international marketers to assess possible cultural differences in values, a methodology and instrumentation is required which permits systematic comparisons of sufficient sensitivity to disclose those differences which might exist. Although there are a variety of segmentation techniques available which the marketer might consider for initially investigating cultural differences, such as benefit segmentation (Haley, 1968), activities, interests and opinions (Wells and Tigert, 1971), life style (Plummer, 1971), motivation research (Pernica, 1974), or cultural assimilators (Fiedler, Mitchell, Triandis, 1971; Worchel and Mitchell, 1972), each has serious limitations. Although all these techniques, except cultural assimilators, have been frequently applied to domestic marketing problems, their application to cross-cultural research is limited. Their conspicuous absence in the sphere of international marketing research stems in large measure from either the vast amounts of data each requires or their difficulty to implement in terms of effort, time and money. For example, to develop a meaningful AIO battery, might require the researcher to generate as many as 300 initial items (Wells and Tigert, 1971). Although cultural assimilators, which are self-administered programmed culture training manuals, provide an apparently effective

method for assisting members of one culture to interact and adjust successfully to members of another, they too require considerable start-up effort. The training manual is composed of a number of "critical incidents" which are elicited from either Americans or foreign nationals. Each incident describes some specific intercultural occurrence or event that made a major difference in each respondent's attitude or behavior toward members of the other culture. The time required to prepare a battery of 75-100 such incidents is approximately 800 man hours (Fiedler, Mitchell, Triandis, 1971).

On the other hand, an approach to segment international markets based on personal value analysis might present fewer limitations. If cultural or subcultural groups were initially found to be dissimilar in their basic value orientations, the marketer might then consider the follow-up use of more complex and costly techniques such as AIO or benefit segmentation. Based on the above, the objectives of this study are primarily four-fold:

- a) to provide empirical evidence that cross-cultural differences in personal value structures do exist.
- b) to investigate the utility of a particularly value assessing instrument, the RVS for disclosing differences in cultural values.
- c) to investigate the extent to which stereotypes of Americans differ across cultures.
- d) to determine the extent to which a person's own value system differs from his stereotype of a person from a different culture.

Personal Values

Domestic Research

The notion that values are of central concern in understanding consumer behavior has been pointed out by a number of researchers (Rosenberg, 1956; Lessig, 1975; Vinson and Munson, 1976; Vinson, Scott and Lamont, 1977). Rosenberg suggests attitudes are derived from values: "the attitude object tends to facilitate the attainment of a number of important values." Similarly Lessig (1975) asserts that attitudes stem in part from the value subsystem and are the expression of values toward given situations and objects. Vinson and Munson (1976) emphasize the centrality of personal values: "they represent the most basic element of the consumer's cognitive world, they structure the individual's perception and understanding of himself, significant others and of the objects and behavior which constitute his psychological environment" (p. 316).

Recent research has shown that not only are values differently held by specific groups of Americans, but more importantly to marketers are findings which suggest values are linked through attitudes to product attributes (Vinson and Munson, 1976; Vinson, Scott and Lamont, 1977; Munson, 1977). Vinson et al (1977) suggest a linkage may exist between general or global values and both more domain specific values relating to specific product qualities as well as subsequent product attribute

evaluations. Vinson and Munson (1976) found parents and their university age children exhibited significantly different preferences for automobile attributes which seemed to be related to their respective underlying value structures. For example, university students attributed significantly greater importance than their parents to the values of "an exciting life" and "pleasure" while their parents placed significantly greater importance on "family security" and being "responsible" (values reflecting standards of social and individual responsibility). Parents emphasized those attributes which signified utilitarian or functional characteristics associated with automobile ownership (e.g., quality of warranty, service required, handling) while students were more concerned with aesthetic and socially observable features (e.g., styling, prestige, luxury interior).

If values serve to segment buyers and are also tied to product attribute evaluations among a rather homogeneous population (e.g., people from the same general culture) one might expect these values to offer even more potential for segmenting peoples of diverse cultural background. This statement however presupposes that the values really do vary significantly across cultures.

Cross Cultural Research

It is a cardinal rule of marketing that it is always easier to appeal to existing cultural wants or expectations than to try to change the culture or to create new needs (Kassarjian and Robertson, 1973, p. 454). If marketers hope to formulate more effective product and promotional strategies for international markets, they must become sensitive to the core values of any given country, and their role in consumption differences. Available research evidence would seem to suggest that success or failure for the international marketer might be related in at least two primary ways to personal values. The first way concerns the importance or centrality ascribed to particular values by another culture, while the second way concerns inappropriate cultural stereotyping of the marketer's culture by the target culture. A few examples will illustrate the types of situations in which a marketer could make the first kind of error. The highly fragmented retail structure in France, with separate shops for fish, meat, pastries, etc., seems to be influenced by traditional French values. The desire to retain family control over the store and maintain status within the community may be more important than showing much profitability. Within American society, the business behavior of many small retailers appears similarly motivated. Wittreich (1962) found owners of small proprietorships were often less concerned with profitability than achieving the feeling of status, prestige, or family security through owning their businesses.

An example of marketing failure related to inappropriate cultural stereotyping is provided by Kitano (1969) in writing about the Japanese-American sub-culture he comments (p. 100), "To many people, Japanese culture suggests mainly tea ceremonies and flower arranging. As often happens, the quaint, the unusual, and even the trivial become so firmly associated in people's minds that complex explanations of behavior are often built upon pursuits that may in fact represent only the interests of a select population." Cultural stereotypes and folklore can be deceiving and perhaps even harmful to the marketer. He should guard against both the possibility of his own inaccurate perceptions of other cultures, or mistakenly assuming the attribution by other cultures of a favorable or specific stereotype of Americans. Chinese visitors to this country are surprised to learn that Americans attribute chop suey to the Chinese, as are Italians upon learning we consider pizza to be an Italian food (Kassarjian and Robertson, 1973).

One approach which might be particularly well suited for assessing values has been developed by Rokeach (1968-69), who views values as having to do with preferable modes of conduct and end-states of existence. The distinction between modes of behavior and end-states of existence implies a differentiation between means and ends or between what Rokeach calls "instrumental" and "terminal" values. Instrumental values relate to modes of conduct and represent a single belief which is personally and socially preferable in all situations with respect to all objects (e.g., ambition, independence and self control). Terminal values are a single belief that some end-state of existence is personally and socially worth striving for (e.g., leading an exciting life, family security and pleasure).

The ability of the Rokeach paradigm to differentiate the value orientations of various domestic groups has been very impressive (Rokeach, 1973). For example, various combinations of terminal and instrumental values have been found which significantly differentiate men from women, "hawks" from "doves," policemen from unemployed blacks (Rokeach, 1968-69). Additionally, some research has been done on the validity and reliability of the RVS itself. Both Rokeach (1973) and Robinson and Shaver (1971) reported relatively high test-retest reliability coefficients for the instrument over three week intervals. Vinson, Munson, and Nakanishi (1977) derived factor structures which clearly indicate the existence of two distinct value categories. A factor analysis of interval scaled importance ratings for each of the items in the RVS yielded 6 factors composed of terminal and 5 factors composed of instrumental values. With only one exception, there was no overlap of instrumental values loading on terminal factors or vice versa.

Hypotheses

It was felt that an adequate demonstration of the utility of the RVS in conducting cross cultural research could be achieved by testing its ability to discriminate between groups which were known to represent differing cultural backgrounds. An attempt was made to include both groups whose cultural backgrounds were assumed to be very different from that of a baseline American group and a second group whose heritage was assumed to be distinct from the American culture, but perhaps somewhat less so. These considerations and the availability of subjects resulted in the selection of Thais and Mexicans. Additionally, evaluations of a common object, the average American stereotype held within each culture, was felt to provide a further test of the sensitivity of the RVS to distinguish perhaps more subtle differences. The following null hypotheses were developed:

- H1: across cultures there will be no differences in the personal value systems of Thais, Mexicans or Americans.
- H2: across cultures there will be no difference in the personal value system each culture attributes to an average American (American Stereotype).
- H3: Within a culture, there will be no difference between that cultures' own value system and the stereotype value system it attributes to an average American (American Stereotype).

While it was strongly anticipated that self value structures might differ markedly between cultures, it was less clear how evaluations of the personal values of the stereotype of an average American might vary. It was felt that cultural norms might serve to not only

identify acceptable behaviors, but also to identify acceptable attitudes toward members of another culture. One might expect the pressure on an individual to be very great to conform to his own cultural's stereotype of another culture. Acting incorrectly, holding the "wrong" stereotype, could meet with social disapproval. One might also anticipate greater homogeneity of the perceptions found within a specific culture than between diverse cultures regarding perceptions of a common third object. Such perceptual homogeneity within culture and heterogeneity across culture could reflect, in part, a) the individual's acceptance of important norms within his own culture, b) cultural perceptual biases and/or c) differences in the information available to members of different cultural backgrounds.

Method

Sample and Instrument

Subjects were drawn from three separate populations. The Thai sample consisted of 32 people selected from a listing of 162 members of the Thai Students' Organization in Los Angeles. A sample of 63 Mexicans was obtained from the members of several Mexican-American organizations in and around the greater Los Angeles area. A sample of 75 Americans was obtained from graduate business students at a large Western university. An attempt to match the samples on major demographics was made. The three samples were very similar with respect to age (25-35 years) and education level (college graduates). A major difference occurred for religion, with the Thai group being 82% Buddhist, while the Mexican group was 85% Catholic, and the American group was primarily Catholic and Protestant.

All respondents completed a questionnaire containing the 36-item RVS and a separate page requesting traditional demographic data. The perceived importance of each value was measured using a seven point Likert-type scale. Each subject provided information on the importance he attributed to each of the 36 values (Self evaluation), as well as his perceptions of the importance placed on each value by an average American (Thai-American stereotype, Mexican-American stereotype and American-American stereotype).

Results and Discussion

Given that one of the main objectives of this study was determination of the capability of personal values as measured by the RVS to reveal cultural differences, stepwise discriminant analysis was employed. The criteria used for controlling the stepwise entry of variables was both the minimum Wilks' lambda and the maximum Mahalanobis distance between groups. The Mahalanobis criterion seeks to maximize the Mahalanobis distance between the two closest groups. The Wilks' lambda criterion is the overall multivariate F ratio for the test of differences among group centroids. The variable which maximizes the F ratio also minimizes Wilks' lambda, a measure of group discrimination.

Ten of the original 36 values were selected using this stepwise procedure (Table 1). The ten values produced a very high degree of separation between the six classification groups as indicated by the final Wilks' lambda (.205), and the canonical correlations for the first five discriminant functions. The square of the canonical correlation indicates that the first discriminant function explains approximately 49%, the second, 30% and the third, 24% of the variance in the grouping variables. Thus the first three functions are highly correlated with the groups while the last two are moderately correlated.

The eigenvalue is a measure of the relative importance of each function. The sum of the eigenvalues is a measure of the total variance existing in the discriminating variables. Expressing each eigenvalue as a percentage of the total sum of eigenvalues indicates the relative importance of the associated function and is shown in the column designated Relative Percentage (Table 1). The Wilks' lambda associated with each function is an inverse measure of the discriminating power in the original variables which has not yet been removed by the discriminant functions. Wilks' lambda was .892 after the 5th discriminant function was derived, which approximates a chi-square of 37.74 ($p < .001$). A lambda of this magnitude or smaller has a $p < .001$ of occurring due to chances of sampling even if there was no further information to be accounted for by the fifth discriminant function.

The relative contributions of the 10 discriminating variables to the respective functions can be seen in Table 2. The first function would seem to primarily represent individual characteristics with an emphasis on the values "broadminded" and "sense of accomplishment" which are over 1.5 times as important as "equality" or "responsibility." F2 is largely composed of "obedience" and "salvation," both making negative contributions. F3 appears most related to "true friendship" and "responsibility," both making negative contributions. Several variables make approximately equal contributions to F4, with "national security" and "responsibility" being the largest negative and "ambitious" the largest positive contribution. F5 is predominately related to "equality."

Discriminant Classification Table

After a discrimination procedure had been established it was of interest to determine whether the discriminant function was useful in classifying people using their scores on the personal values into their own cultural group. The classification matrix (Table 3) provides a convenient method of summarizing the number of correct and incorrect classifications made by the discriminant function. The diagonal elements denote the number of correct classifications (hits) and the off-diagonal elements denote the number of incorrect classifications (misses).

The classification power of the discriminant procedure is quite encouraging in that 65% of the cases are correctly identified as compared to an expected 22% if group assignment were to the most frequent category (i.e., the maximum chance criterion). Morrison (1969) argues in favor of a more lenient chance model which he calls the proportional chance criterion. Under this latter chance model, the chance criterion is adjusted to take into account the fact that some classifications are being made to the smaller groups. In our case, this proportional chance criterion would yield a 13% chance hit rate. However, we utilize the more conservative criterion of 22% in light of the fact that the Thai sample size was judged to be insufficient to permit the use of both an estimation and validation sample. To statistically test whether the 65% hit rate was better than the 22% maximum chance criterion a Q statistic was calculated (Press, 1972). The Q value exceeded a value of $X^2 = 10.82$ ($p < .001$), hence the classification procedure does significantly better than chance.

Hypothesis 1, 2 and 3 regarding the differences in the self value and stereotype value structures between cultures were tested by comparing the centroids derived for each group in the five dimensional discriminant space. The centroids represent mean discriminant scores for each group on the respective discriminant functions.

TABLE 1

Discriminant Analysis Summary Table

Step Number	Variable Removed	F to Enter or Remove	Wilks' Lambda	Significance
1	Obedient	23.12	.742	.0
2	Responsible	21.18	.563	.0
3	Cheerful	11.09	.482	.0
4	Ambitious	6.92	.437	.0
5	National Security	6.81	.396	.0
6	Broadminded	14.92	.323	.0
7	Equality	8.17	.287	.0
8	Sense of Accomplishment	10.28	.248	.0
9	Salvation	4.14	.233	.0
10	True Friendship	8.90	.205	.0

Discriminant Function	Eigenvalue	Relative Percentage	Canonical Correlation	Functions Derived	Wilks' Lambda	Chi-Square	DF	Significance
				0	.2052	524.17	50	.0
1	.9250	46.44	.693	1	.3951	307.38	36	.0
2	.4321	21.70	.549	2	.5658	188.49	24	.0
3	.3131	15.72	.488	3	.7430	98.32	14	.000
4	.2008	10.08	.409	4	.8922	37.74	6	.000
5	.1207	6.06	.328					

TABLE 2

Standardized Discriminant Function Coefficients

	F1	F2	F3	F4	F5
Equality	.260	.020	-.064	.158	-1.115
National Security	-.088	.096	.289	-.481	.179
Salvation	.178	-.332	.124	.392	.055
True Friendship	-.202	.188	-.663	.088	.336
Sense of Accomplishment	.414	.097	.250	.255	.226
Ambitious	.123	.295	.077	.450	.247
Broadminded	.435	.056	.214	.273	.542
Cheerful	.170	.126	.107	.378	.112
Obedient	.288	-.743	.315	.139	.151
Responsible	.231	.345	-.470	-.414	.227

The F ratios based on the Mahalanobis distances between all pairs of group centroids were significant beyond the $p < .001$ level. Therefore one can reject the hypotheses that the value profiles from any two groups are equal. The significance of these findings suggest at least the following conclusions:

1. Self value systems do differ markedly across the three cultures studied.
2. Conceptions of American stereotypes also differ markedly across cultures. Apparently there is not one universally held American stereotype.
3. Within cultures there are significant differences between self value systems and those values attributed to an American stereotype.

Given that the primary thrust of this research was on the ability of personal values to discriminate various cultural groups, and the applicability of the RVS to this task, it was not the intent to focus on differences in specific isolated value items. However, readers interested in making specific comparisons are referred to Table 4 which presents mean importance ratings on both self value and American stereotype value structures for each culture. The univariate F ratios indicate that at least one pair of the groups are significantly different on each value. Some interesting differences in the evaluations of both instrumental and terminal values are observable across the three groups. For

example, This rate the terminal values of "a comfortable life" and an "exciting life" as significantly less important than either Americans or Mexicans. However they attach significantly greater importance than Americans to those terminal values related to social harmony - that is, "world at peace," "national security" and "family security." On the other hand, This attach significantly less importance than Americans to the instrumental values of "ambition" and "imagination."

Conclusion

Personal values represent the most fundamental element of the consumer's cognitive world. They enable the individual to structure his perception and understanding of himself, of significant others and of the objects embodied within his psychological reality. Inasmuch as it has been repeatedly observed that values are antecedent to many types of behavior, one might expect differential value orientations to manifest themselves in differential overt behavior. Knowledge of the value orientations held by a specific group may help in understanding, explaining, or perhaps predicting subsequent attitudes and behavior. However there is a limited empirical research, outside the current study, which focuses on either value measurement or their potential usefulness to those engaged in international marketing.

The findings of this study indicate that personal values can be productively employed to discriminate peoples

TABLE 3
Classification Matrix

Actual Group ^a	No. of Cases	Predicted Group					
		Thai Self	Thai American	Mexican Self	Mexican American	American Self	American Stereo
Thai Self	32	18 56.3%	2 6.3%	6 18.8%	3 9.4%	2 6.3%	1 3.1%
Thai American	32	6 18.8%	16 50.0%	3 9.4%	2 6.3%	4 12.5%	1 3.1%
Mexican Self	63	3 4.8%	4 6.3%	46 73.0%	2 3.2%	6 9.5%	2 3.2%
Mexican American	63	1 1.6%	2 3.2%	19 30.2%	30 47.6%	3 4.8%	8 12.7%
American Self	75	5 6.7%	3 4.0%	4 5.3%	8 10.7%	50 66.7%	5 6.7%
American Stereo	75	0 0%	0 0%	0 0%	4 5.3%	10 13.3%	61 81.3%

Percentage of Grouped Cases Correctly Classified: 65.0%

^aThai Self, Mexican Self and American Self are the Self evaluations derived from each respective culture. Thai American, Mexican American and American Stereo are the stereotypes of an average American held by each respective culture.

with culturally diverse backgrounds. Significant differences in the self value orientations were observed between three cultural groups: Thais, Mexicans, and Americans. Significant differences were also found between the three groups in the stereotype each culture holds of an average American's personal value system. Additionally, within each culture significant differences were observed between a person's self values and those he attributed to an average American. Discriminant functions derived from responses to the Rokeach Value Survey were able to correctly classify 65% of the people according to their cultural background, significantly better than the 22% expected by chance.

Because of the robust ability of values to distinguish among peoples of different culture, value analysis may represent a convenient and economical technique to aid in market segmentation strategies. The RVS might offer marketers and researchers a valuable tool for identifying normatively based cross cultural differences. It is easily administered and could be employed to develop a profile of a general population or profiles of individual segments prior to the application of other more expensive and complex segmentation techniques. Value analysis should become an increasingly important area of research for the international marketer.

TABLE 4

Summary of Mean Importance Value Ratings
for Thais, Mexicans, Americans and American
Stereotypes with Univariate F-Tests

<u>Terminal Values</u>	<u>Thai Self</u>	<u>Thai American</u>	<u>Mexican Self</u>	<u>Mexican American</u>	<u>American Self</u>	<u>American Stereo</u>	<u>F-Ratio^a (df=5,334)</u>
Comfortable Life	4.78	5.50	5.82	5.25	5.40	5.80	6.83
Exciting Life	4.84	6.06	6.00	5.95	5.34	4.61	18.01
World at Peace	6.34	5.09	6.00	5.53	5.13	5.02	8.93
Equality	5.96	5.93	6.42	4.61	4.93	4.33	25.72
Freedom	6.59	6.75	6.50	5.88	6.08	5.56	10.46
Happiness	6.53	6.06	6.55	5.33	5.73	5.65	10.00
National Security	6.18	5.56	5.46	6.03	4.81	5.30	7.65
Pleasure	5.37	6.31	6.06	6.34	5.52	5.61	8.19
Salvation	5.75	4.06	5.34	5.19	3.66	3.94	15.63
Social Recognition	4.62	4.50	5.52	6.22	4.90	4.76	14.70
True Friendship	6.21	4.71	6.23	6.26	5.96	5.10	14.30
Wisdom	6.37	5.50	6.14	5.66	5.32	3.70	36.00
World of Beauty	4.87	5.28	5.58	5.28	4.90	3.41	19.63
Family Security	6.68	4.78	6.50	6.34	5.69	5.18	14.49
Mature Love	6.37	3.34	6.49	5.63	5.78	4.74	36.49
Self Respect	6.75	6.71	6.60	5.69	6.22	5.29	14.40
Sense of Accomplishment	6.68	6.34	6.12	5.79	5.30	4.30	24.59
Inner Harmony	6.34	4.68	6.33	5.26	5.60	4.46	18.90
<u>Instrumental Values</u>							
Ambitious	4.65	5.34	5.77	6.15	5.33	4.94	8.69
Broadminded	6.65	6.12	6.09	5.07	5.37	3.74	32.58
Capable	6.34	6.18	5.96	5.80	5.65	4.85	14.12
Cheerful	5.06	5.37	6.03	5.25	4.93	4.29	16.13
Clean	6.31	5.21	5.79	5.30	4.22	4.36	20.61
Courageous	5.53	5.00	5.95	5.34	5.33	4.88	4.36
Forgiving	5.06	4.87	5.76	4.80	5.08	4.00	10.81
Helpful	5.34	4.09	5.84	4.61	4.77	4.12	13.40
Honest	6.34	5.28	6.38	4.96	5.97	4.86	15.93
Imaginative	4.15	5.03	5.93	5.39	5.20	3.70	21.10
Independent	5.37	6.59	6.11	5.85	5.72	4.89	13.97
Intelligent	6.53	5.68	5.92	5.58	5.37	3.80	32.72
Logical	6.34	5.59	5.68	5.53	5.30	3.97	18.66
Loving	5.93	4.65	6.06	5.14	5.50	4.73	11.23
Obedient	4.56	2.62	5.49	5.11	3.81	4.10	23.12
Polite	5.71	3.96	5.69	5.03	4.98	4.29	14.23
Responsible	6.81	5.37	6.30	5.42	5.85	4.29	20.93
Self-Controlled	5.37	4.78	5.85	5.15	5.09	4.38	12.24

^aAll tabled F ratios are significant at $p < .01$. F critical (5,334,.01) = 3.03.

References

- Fred E. Fiedler, Terence R. Mitchell and Harry C. Triandis, "The Culture Assimilator: An Approach to Cross Cultural Training," Journal of Applied Psychology, Vol. 55, No. 2 (1971), 95-102.
- Russell J. Haley, "Benefit Segmentation: A Decision-Oriented Research Tool," Journal of Marketing, (July 1968), 30-35.
- Walter A. Henry, "Cultural Values Do Correlate with Consumer Discontent," Journal of Marketing Research, 13 (May 1976), 121-127.
- Harold H. Kassirjian and Thomas S. Robertson, Perspectives in Consumer Behavior, (Glenview, Illinois: Scott Foresman and Co., 1973).
- Harry H. L. Kitano, Japanese Americans, (Prentice-Hall, 1969).
- P. V. Lessig, "A Measurement of Dependencies Between Values and Other Levels of the Consumer's Belief Space," Journal of Business Research, 3 (July 1975), 227-241.
- Donald G. Morrison, "On the Interpretation of Discriminant Analysis," Journal of Marketing Research, 6 (May 1969), 156-163.
- J. Michael Munson, "An Investigation of the Rokeach Value Survey for Applications to Cross Cultural Research," in J. Stolen and J. Conway, eds., Proceedings of the American Institute of Decision Sciences, 1977.
- Joseph Pernica, "The Second Generation of Market Segmentation Studies: An Audit of Buying Motivation," in W. D. Wells, ed., Life Style and Psychographics, (Chicago: American Marketing Association, 1974), 277-313.
- Joseph T. Plummer, "Life Style Patterns and Commercial Bank Credit Card Usage," Journal of Marketing, (April 1971), 35-41.

S. James Press, Applied Multivariate Analysis, (New York: Holt Rhinehart and Winston, 1972).

J. P. Robinson and P. R. Shaver, eds., Measures of Social Psychological Attitudes, (Ann Arbor: Institute for Social Research, 1971), 463.

Milton J. Rokeach, "The Role of Values in Public Opinion Research," Public Opinion Quarterly, (Winter 1968-69), 547-549.

Milton J. Rokeach, The Nature of Human Values, (New York: The Free Press, 1973).

Milton J. Rosenberg, "Cognitive Structure and Attitudinal Affect," Journal of Abnormal and Social Psychology, 53 (November 1956), 367-72.

J. E. Scott and L. H. Lamont, "Relating Consumer Values to Consumer Behavior: A Model and Method for Investigation," in Thomas V. Greer, ed., Increasing Marketing Productivity, (Chicago: American Marketing Association), 283-288.

Donald E. Vinson and J. Michael Munson, "Personal Values: An Approach to Market Segmentation," in K. L. Bernhardt, ed., Proceedings, Marketing: 1776-1976 and Beyond, (Chicago: American Marketing Association, 1976), 313-317.

Donald E. Vinson, J. Michael Munson and Masao Nakanishi, "An Investigation of the Rokeach Value Survey for Consumer Research Applications," in W. O. Perreault, ed., Proceedings, Association for Consumer Research, 4 (1977), 247-252.

Donald E. Vinson, Jerome E. Scott and Lawrence M. Lamont, "The Role of Personal Values in Marketing and Consumer Behavior," Journal of Marketing, 41 (April 1977), 44-50.

William D. Wells and Douglas J. Tigert, "Activities, Interests and Opinions," Journal of Advertising Research, (August 1971), 27-34.

Warren J. Wittreich, "Misunderstanding the Retailer," Harvard Business Review, 40 (May-June 1962), 147-159.

Steven Worchel and Terence R. Mitchell, "An Evaluation of the Effectiveness of the Culture Assimilator in Thailand and Greece," Journal of Applied Psychology, Vol. 56, No. 6, (1972), 472-479.

R. Neil Maddox, University of Missouri-St. Louis
 Kjell Grønhaug, Norwegian School of Economics and Business Administration
 Richard E. Homans, University of Missouri-St. Louis
 Frederick E. May, University of Missouri-St. Louis

Abstract

Interviews conducted with new automobile purchasers in the U.S. and Norway are analyzed in this study. Correlates of information gathering activities and the size of the evoked set in the two samples are compared. Similar patterns of relationships are found in the two samples. The Norwegian respondents generally reported higher levels of search which seems consistent with known structural differences between the two cultures.

Introduction

Even though the influence of culture upon consumer behavior is widely acknowledged, cross-cultural studies of the consumer's decision process are conspicuous by their absence (Nicosia and Mayer, 1976). While there are notable exceptions (Douglas, 1976; Pfaff and Blivice, 1977), the costs and difficulties of cross-cultural research seem to have dissuaded most consumer behavior researchers from initiating investigations spanning national boundaries (Sheth, 1974).

Both theoretical and managerial concerns dictate pursuing studies in a number of different settings. From a scholarly point of view, such studies offer an invaluable arena for developing and testing theories. It is frequently held that models of consumer behavior capture fundamental human decision processes (Markin, 1974, p. 55). If so, these models should be applicable in a wide range of situations if the background variables are known. Such beliefs cannot be tested if empirical studies are only conducted within the relatively homogeneous confines of a single society. Our confidence in the generality of a theory is heightened if its descriptions of a process are found to be accurate in a wide range of settings.

More pragmatically, given the rising importance of multinational businesses, cross-cultural studies are vitally important to the marketing manager. Many foreign markets are rapidly developing from a state of scarcity to one of abundance (Sheth, 1974). Consumers in a growing number of countries have a wide range of choices among products and brands. They have, perhaps for the first time, the opportunity to make consumption decisions. Just as was the case in the United States (Engle, Kollat and Blackwell, 1973, p. 46), as goods become plentiful it becomes imperative to develop an understanding of the factors influencing consumers as they make their choices.

Further, efficiency in the employment of marketing resources requires the identification of similarities as well as differences between cultures (Douglas, 1976). If the multinational firm is to succeed in truly developed markets, it must realize economies of scale in the application of its marketing efforts that are not available to its national or regional competitors. If the multinational corporation serves only an *ad hoc* collection of discrete national markets that must be approached independently, it will be at a disadvantage. Without some offsetting competitive advantage, the larger firm merely adds layers of overhead costs that must eventually be borne by each of the local operations. If it is to become a truly viable institution in a world of developed economies, the multinational firm must be able to serve individual markets at least as efficiently as its nation-

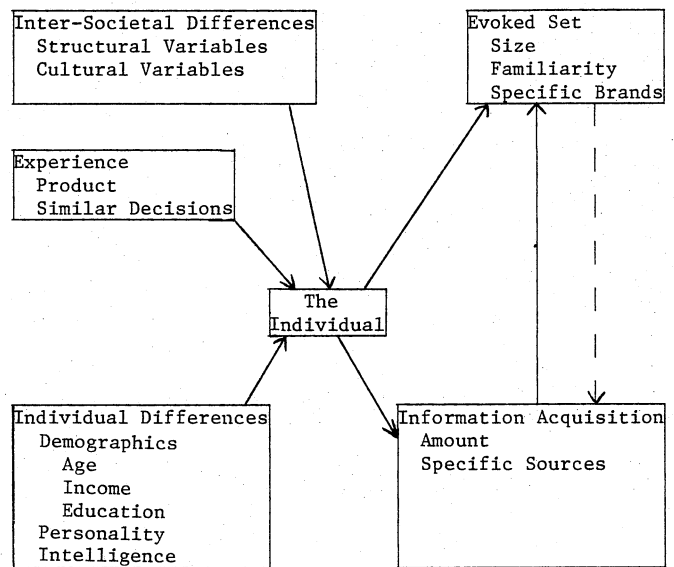
al counterparts. It would be desirable to identify cross-national similarities that can be used to segment markets and develop common marketing strategies.

Detailed explorations of cross-cultural similarities in consumer decision processes will entail large-scale and expensive research efforts. Before considering or recommending such an undertaking, one should wish to see some indications of the likelihood of its success. This paper represents a modest attempt at such a demonstration. Converging findings from studies of automobile decision making in two different cultures are discussed.

A Simplified Model

Extensive models of the consumer decision process have been presented by a number of authors (Nicosia, 1966; Howard and Sheth, 1969; Engel, Kollat and Blackwell, 1973). These are overly elaborate for the present purposes, given the state of the art in cross-cultural studies of consumer behavior. We will constrain the discussion to the simpler model shown in Figure 1. At the beginning of any decision problem the consumer has stored in his mind a bundle of attitudes, beliefs, knowledge, evaluative criteria, etc. These result from the unique characteristics of the society within which he lives, individual differences among the members of that society (e.g., personality, intelligence, and demographics), his previous experience with the product and the making of similar decisions.

FIGURE 1
 A Simplified Model of the Consumer Decision Process



Societal differences are viewed as being of two sorts. First, there are the more measurable structural differences such as income levels, the general development of the economy and the availability and cost of the product and product related information. Secondly, there are the more ephemeral aspects of culture including norms, values and mores, etc.

These exogenous or background variables are fixed at the onset of the decision task. They are shown as influencing both the acquisition of information and the formation of the evoked set, that limited group of alternatives actively considered by the decision maker (Howard and Sheth, 1969, p. 26). Information gathering may vary both in terms of the overall level of activity and the consultation of specific sources. Evoked sets may be analyzed in terms of their size and the balance between familiar and unfamiliar brands as well as according to the specific brands which they contain.

Information seeking should influence the size and composition of the evoked set. Finally, a feedback relationship is shown, since the perceived satisfactoriness of the evoked set would be expected to affect the course and duration of information seeking activities.

Studies of these relationships in distinctly different cultural settings are discussed in the following sections. It should be emphasized that our concern does not lie with demonstrating the presence or absence of between-culture differences with respect to the level on any variable. Rather, we are interested in exploring cross-cultural similarities in the relationships between the different sets of variables.

The Studies

Two nearly contemporaneous studies are compared and contrasted here. The United States study was conducted in St. Louis, Missouri. Data for the Norwegian study were gathered in Bergen, Norway.

Samples

Norwegian Study. Personal interviews averaging 45-60 minutes were conducted with a sampling of new car registrants during the late fall of 1969. Time and financial constraints forced limitation of the sample to Bergen, Norway's second largest town. Individuals were selected from transcripts of new car registrations. After as many as four recalls, completed questionnaires were obtained from 101 of the 138 individuals drawn for the sample. Five of the interviews could not be used because the automobiles were found to have been purchased for some commercial purpose or due to incomplete data.

United States Study. The data for this study are from personal interviews with 387 household heads residing in the St. Louis metropolitan area. A standard cluster-area probability sampling procedure was used to select a sample of dwelling units in the metropolitan area. Interviews were conducted by the Public Opinion Survey Unit of the University of Missouri in the Spring of 1966. Up to four interviewer calls were made at each selected dwelling unit. The response rate of 77 percent of attempted interviews appears reasonable for a large metropolitan area and for a high proportion of nighttime interviews. Respondents were asked to recall and describe all cars owned and previously owned by themselves and their family. Next, each respondent was asked to describe the decision leading to his or her most recent purchase. The analysis reported here is based upon data from those 132 household heads whose most recent purchase was a new car, who owned at least one car before the most recent purchase, whose brand choice was not restricted by special deals, and whose answers to the questions used were ascertained.

Measures and Analysis

Each study gathered a large amount of information about the individual, his automobile ownership history and his most current auto purchase decision. A number of questions reflecting on the relationships shown in Figure 1 were common to both studies. These are reproduced in Appendix A.

It seemed desirable to make minimal assumptions about the level of measurement. Even though some variables were measured on an interval scale, Kendall's Tau, which requires only ordinal measures, was used throughout the analysis unless another statistic is indicated.

Findings and Discussion

The patterns of findings shown in Tables 1-4 are supportive of the notion that similar processes operate in the two cultural settings. Where differences are noted at least some explanation can be offered in terms of the effects of structural differences.

Demographic Variables and Evoked Set Size

The relationships between demographic variables and the size of the evoked set are shown in Table 1. The similarities between the two studies are striking. Age and education are significantly related to evoked set size in both studies; income is found to be lacking in explanatory power. The direction of the significant relationships is as expected. As the individual ages, he accumulates experience with the product and product related decisions and he develops a more completely formed preference ordering (Howard and Sheth, 1969, pp. 26-29). The elimination of marginally interesting brands, coupled with a finite upper limit on the number of brands of automobiles available, would lead to smaller evoked sets. Similarly, a positive relationship between evoked set sizes and education was anticipated. If the educational system is successful, it should develop an individual with heightened capabilities and inclinations to retain and manipulate information. An educated person would be expected to discover and consider more alternatives than one not so favored.

TABLE 1
Relationships Between Evoked Set Size
And Demographic Variables^a

	Norwegian Study	U.S. Study
Age	-.12*	-.09*
Income	-.05	.05
Education	.22***	.16***

^aKendall's Tau

*p ≤ .10, one-tailed test

**p ≤ .05, one-tailed test

***p ≤ .01, one-tailed test

Demographic Variables and Information Seeking

These relationships are shown in Table 2. The pattern is much the same as that noted in Table 1. Age shows significant negative correlations with all measures of information seeking. Similar arguments to those advanced in the preceding paragraphs apply here; as the individual develops greater total experience he has less need to gather information specifically for the present decision.

With one nonsignificant exception, the relationships with education are, predictably, positive. The fact that a far greater number of significant correlations are found in the Norwegian study does not admit to simple explanations. Perhaps this reflects the emphasis upon consumer education found in Scandinavian educational systems (Thorelli, 1977).

The findings as to the effects of income seem confusing. The Norwegian data show consistently negative signs. This would support a hypothesis that the higher the income the lower the perceived risk and, hence, the need to acquire information as a risk reducing mechanism (Engel, Kollat and Blackwell, 1973, p. 587). This hypothesis would not explain the United States data where the sig-

nificant relationships are positive. In this instance availability may be the key. The significance of advice results from the tendency of higher income persons to consult with individuals having technical or professional expertise; income is negatively related to discussing the purchase with friends, relatives and co-workers (data not shown). Higher income persons may have easier access to those with professional knowledge of automobiles. A similar situation may prevail with brochures. A substantial number of those reporting they had read brochures indicated that these had been mailed to them by the seller without any initiative on the buyer's part. Since they are attractive prospects, this may reflect the fact that higher income consumers are likely to find themselves on a large number of mailing lists. The interesting thing is that these unsolicited brochures were not only received, but also read and remembered.

TABLE 2
Relationships Between Demographic Variables^a
And Use Of Information

Information Source	Age		Income		Education	
	Norway	U.S.	Norway	U.S.	Norway	U.S.
Advice Seeking	-.16**	-.11*	-.02	.15**	.15*	.14**
Reading Advertisements	-.10	-.12**	.07	-.04	.08	.06
Reading Brochures	-.12*	-.00	-.01	.12*	.18**	.02
Dealer Visits	-.11	-.17**	-.05	.05	.22***	.09
Test Driving	-.12*	-.12*	-.14*	.07	.24***	-.01
Total Sources	-.12*	-.19***	-.07	.14**	.17**	.16**

^aKendall's Tau
*p ≤ .10, one-tailed test
**p ≤ .05, one-tailed test
***p ≤ .01, one-tailed test

The (nonsignificant) negative sign found for test driving in the United States data may be reflecting the influence of compliance rather than education. Paraphrasing one respondent, "The salesman is supposed to get you to test it, so you drive it a few blocks so he doesn't get in trouble." Perhaps less educated respondents are more easily cowed by auto salesmen.

Information Seeking and Evoked Set Size

It was expected that the level of information seeking would be positively related to the size of the evoked set. Two considerations gave rise to this prediction. First, both the level of information seeking and the evoked set are believed to be associated with the complexity of the decision process. Small evoked sets (approaching one) and minimal information gathering are characteristic of routine response behavior (Howard and Sheth, 1969, p. 528). Problem solving (limited or extended) is believed to be associated with larger evoked sets and more active efforts to acquire information.

As shown in Table 3, this expectation is supported by both sets of findings: larger evoked sets are positively related to more intense information gathering. With the exception of advice, stronger relationships are found for the Norwegian respondents. This can, perhaps, be explained by structural differences between the two societies. At the time of the studies automobiles were two or three times as expensive in Norway as in the United States and incomes were roughly 30-40% lower. As a result, a far smaller proportion of the households owned a car. Since the Norwegian buyer was committing a larger proportion of his income to a product with which he was less familiar than his United States counterpart, he

would be expected to engage in more extensive problem solving.

TABLE 3
Relationships Between Evoked Set Size^a
And Use Of Information

Information Source	Norwegian Study	U.S. Study
Advice Seeking	.13*	.14**
Reading Advertisements	.21***	.08
Reading Brochures	.39***	.00
Dealer Visits	.62***	.11*
Test Driving	.42***	.08
Total Sources	.43***	.19***

^aKendall's Tau
*p ≤ .10, one-tailed test
**p ≤ .05, one-tailed test
***p ≤ .01, one-tailed test

Summary

The case for cross-national studies of consumer behavior was argued in the introduction. Such studies are needed both for theory development and for the managerially useful insights which may be gained. A simplified model was then presented. Findings from studies of Norwegian and American automobile purchasers on the relationships contained in this model were then compared. Differences in the level or intensity of information search activities were noted. It seemed that these differences could be attributed to structural differences between the two cultures. However, the most notable findings were the similarities in the nature of the relationships between variables. There seemed to be strong indications that similar underlying processes were operating in these quite different societies.

Appendix Questions Asked of Respondents and Their Coding

	Norwegian Study	U.S. Study
Advice	"Did you seek any advice from others when you were buying the car?" ¹	"Did you hear anyone talk about this make of car or ask anybody's advice before you bought it?" ¹
Brochures	"Did you study any brochures before purchase?" If yes. "For which car(s)?" ¹	"Did someone selling the car call or write or give you some literature...?" "Did you, yourself, go out to look at the car, or get some literature...?" ¹
Dealer Visit	"Did you visit any dealer(s)?" If yes. "Whom?" ³	"Did you, yourself, go out to look at the car, or get some literature or talk to someone who was trying to sell it?" If yes. "Was it just one contact, or were there several?" ^{1,2}
Test Driving	"Did you test drive any car(s) before purchase?" If yes. "Which ones?" ³	"Did you drive this car or a similar model before you bought it?" ¹

Appendix (con't.)

	Norwegian Study	U.S. Study
Advertisements	"Did you study any advertisements for any car before purchase?" If yes. "For which cars?" ¹	"Did you hear or see anything about this car on television or radio, in a newspaper, magazine, or buying guide?" If yes. "Was it an advertisement, or a story that you remember?" ¹
Total	Sum of sources coded above. ³	Sum of sources coded above. ³
Income	"As an estimate, what was your family's income last year before taxes?" ³	"Which of these categories best describes your total family income before taxes last year?" ³
Education	"Did you continue your education after public school?" If yes. "How many years?" ³	"How many grades of school did you finish?" ³

¹Dichotomous coding: named --- did not name this source

²Coded as only one visit --- more than one visit

³Coded as interval data

References

Susan P. Douglas, "Cross National Comparisons and Consumer Stereo Types: A Case Study of Working and Non-Working Wives in the U.S. and France," Journal of Consumer Research, 3(June, 1976), 12-20.

James F. Engel, D. T. Kollat and R. D. Blackwell, Consumer Behavior, Rev. Ed. (Homewood, Ill.: Dryden Press, 1973).

John A. Howard and J. N. Sheth, The Theory of Buyer Behavior (New York: John Wiley & Sons, Inc., 1969).

Rom J. Markin, Jr., Consumer Behavior: A Cognitive Orientation (New York: MacMillan, 1974).

F. M. Nicosia, Consumer Decision Processes: Marketing and Advertising Implications (Englewood Cliffs, N.J.: Prentice-Hall, 1969).

Francesco M. Nicosia and R. N. Mayer, "Toward a Sociology of Consumption," Journal of Consumer Research, 3(September, 1976), 65-75.

N. H. Nie, C. H. Hull, J. G. Jenkins, K. Steinbrenner and D. H. Bent, Statistical Package for the Social Sciences, 2nd Ed. (New York: McGraw-Hill Book Company, 1975).

Martin Pfaff and S. Blivice, "Socio-Economic Correlates of Consumer and Citizen Dissatisfaction and Activism," Paper presented to The Research Symposium on Consumer Satisfaction/Dissatisfaction and Complaining Behavior, Indiana University, Bloomington, Ind., April 21-22, 1977.

Jasdish N. Sheth, "The Next Decade of Buyer Behavior Theory and Research," in J. N. Sheth, Ed., Models of Buyer Behavior: Conceptual, Quantitative and Empirical (New York: Harper and Row, 1974).

Hans G. Thorelli and Y. R. Puri, "On Complaining in Norway and the Role of Information Seekers," Paper presented to The Research Symposium on Consumer Satisfaction/Dissatisfaction and Complaining Behavior, Indiana University, Bloomington, Ind.: April 21-22, 1977.

SELECTING PARETO OPTIMAL SUBSETS FROM MULTIATTRIBUTE ALTERNATIVES

James B. Wiley
Wayne State University and University of Florida

Abstract

Previously, Green (1974) illustrated the use of fractional replication procedures for selecting multiattribute alternatives for conjoint measurement studies of consumer choice. This study suggests the information laden choices in conjoint measurement are among alternatives in pareto optimal subsets. Pareto optimal subsets are those in which no alternative "dominates" another on every attribute. Procedures for selecting pareto optimal subsets are illustrated.

Introduction

How do consumers choose among multiattribute alternatives? Recent studies by Green and Rao (1972), Green, Carmone, and Wind (1972), Fiedler (1972), and Bettman, Capon, and Lutz (1973, 1974) have investigated this question using the techniques of conjoint measurement (Krantz and Tversky, 1972), trade-off analysis (Johnson, 1974), or functional measurement (Anderson, 1972). At the core of the many similarities among these three techniques are the following four fundamental premises: a) individuals (or other decision-making units) assign "partworths" to attribute levels, b) "utility" is a function of these partworths, c) choice is proportional to utility, and d) although the individual may be unable to articulate the partworths, they may be inferred from numerous choices among alternatives having characteristics which are varied in systematic ways. In this paper we use "conjoint measurement" in a generic sense, to refer to the general problem of inferring "partworth's" from choices.

Conjoint measurement is a recent development in mathematical psychology that is concerned with resolution of the joint effect of a set of attributes (measured at the ordinal level) into separate, independent contributions, which are measured at the level (asymptotically) of interval scales with common unit. That is, it is concerned with the effect of two or more independent variables on the ordering of a dependent variable. The approach is designed to solve a fundamental scientific problem, that of determining the rules by which a set of independent variables can be combined and used to predict values for a dependent variable. In the behavioral sciences, however, only rarely can one measure the dependent variable at the interval level, let alone the independent variables. In order to bridge this discrepancy between technical feasibility and scientific objective, conjoint measurement is used to determine whether there exists any rescaling of the dependent and independent variables, that when combined according to a hypothesized composition rule, would yield results consistent with the observed ordinal data structure. Alternatively, conjoint measurement may be used as a scaling method. When used as a scaling method, one tries to solve the measurement and composition rule problems together by fitting scales that obey the stated composition rule to some suitable approximation (Green and Wind, 1973).

With the additive conjoint measurement algorithms, such as Kruskal's MONANOVA (1965), an additive composition rule is assumed. Then, the above four premises can be represented as follows. The utility $U(x)$ of some multiattribute alternative $x = (x_1, x_2, \dots, x_n)$,

expressed as an n -component vector, is defined as,

$$U(x) = f(x) = \sum_{j=1}^n u_j(x_j) \quad (1)$$

where u_j is a real-valued function of x_j , and x_j is the "level" of alternative x on the j (th) attribute.

As Dawes and Corrigan (1973) point out, additive models typically are applied to problems in which it also is reasonable to assume that the real-valued function u_j (1) is conditionally monotonic over x_j (which also implies an a priori ordering over x_j .) That is, given two alternatives, $A = (x_1, x_2)$ and $B = (x_1, x_2')$, then

$$U(A) \leq U(B) \text{ iff } (x_1, x_2) \leq (x_1, x_2') \quad (2)$$

which implies utility is monotonic with x_2 and, in general, with x_i . As a consequence, we can speak of an alternative, A , "dominating another alternative, B , whenever

$$x_i > x_i' \cap x_j \geq x_j' \cup (x_i \geq x_i' \cap x_j > x_j') \quad (3)$$

Under such conditions, choice of $A = (x_i, x_j)$ must obtain by virtue of the assumptions of the model and a priori information regarding the setting. Since an alternative that is equal to or greater than another alternative on every attribute must be chosen by assumption, it follows that valuable information for scaling purposes (as distinct from model testing) is gleaned from choices among sets in which no alternative "dominates" another. Alternatives in such sets are said to be Pareto optimal. This paper describes how to select Pareto optimal subsets of multiattribute alternatives constructed from p^p factorial series. This series includes three attributes at three levels designs, four attributes at four levels, five attributes at five levels and so forth. Pareto optimal subsets for the 3^3 through the 6^6 series are reported. A desirable property of the reported sets is that all subsets selected from the reported sets will also be Pareto optimal. As a consequence, Pareto optimal subsets can be organized according to balanced incomplete block designs (BIBD's), with the alternatives within blocks Pareto-optimal.

An Example

Suppose we wish to determine the relative importance of transportation characteristics. Conversations with experts and traffic managers as well as a review of the literature suggests three dominant service criteria; reliability, rate, and transit time. Further conversation with experts suggest the following levels are realistic for a given origin/destination/commodity shipment:

Reliability	on days late			
	time	1	2	3
A	95%	3%		
B	88%	8%	2%	
C	80%	12%	4%	2%

Rate	
A	\$1.13/cwt.
B	1.54/cwt.
C	1.95/cwt.

Normal Transit Time	
A	2 days
B	4 days
C	6 days

For example, one alternative may cost \$1.13 per hundred weight, take six days normal transit time, with an expectancy that 95% of the time it will arrive on or before the sixth day and a 3% chance it will arrive on the seventh day. There are 27 such alternatives that can be constructed from this 3³ design.

Aside from the so-called "alphabet routes" which it is said once were favored for some commodity shipments, it is reasonable to assume that a decision-maker would prefer a cheaper, quicker, more reliable alternative to a more expensive, slower, less reliable alternative. Informative choices, therefore, are among alternatives that require the decision-maker to forego performance according to one criteria in order to gain performance according to another criteria. Requiring the decision-maker to rank order the following alternatives requires such choices.

A PARETO OPTIMAL SUBSET OF 3³ DESIGN

Alternative	Reliability	Rating	Transit Time
1	(A) 95% 3%	(B) \$1.54/cwt.	(C) 6 days
2	(A) 95% 3%	(C) 1.95/cwt.	(B) 4 "
3	(B) 88% 8% 2%	(A) 1.13/cwt.	(C) 6 "
4	(B) 88% 8% 2%	(C) 1.95/cwt.	(A) 2 "
5	(C) 80% 12% 4% 2%	(A) 1.13/cwt.	(B) 4 "
6	(C) 80% 12% 4% 2%	(B) 1.54/cwt.	(A) 2 "

While the above list is not exhaustive (for example, the alternative B(2), B(2), B(2) could be added, the alternatives do comprise a balanced, pareto optimal set which can be constructed following a simple heuristic.

How to Construct Balanced Pareto Optimal Subsets

A simple procedure for constructing a balanced set of Pareto optimal alternatives is based on cyclic hyper-graeco-latin square designs. This somewhat awesome term describes the straightforward generalization of latin square designs with multiple factors. For example, a cyclic graeco-latin square of a 3³ design can be represented as:

		COLUMNS		
ROWS	1	2 3	2 3 1	3 1 2
	2	3 1	3 1 2	1 2 3
	3	1 2	1 2 3	2 3 1

The position in a given cell refers to attribute and number in that position refers to attribute level. Three alternatives are represented in a given row or column and these three alternatives are Pareto optimal. To return to the above example, the first row represents alternatives 1, 4 and 5 respectively of Table 1.

The above design can be represented as a composite of the following three latin squares:

TABLE TWO
ELEMENTS OF COMPOSITE DESIGN

(1)	(2)	(3)
1 2 3	2 3 1	3 1 2
2 3 1	3 1 2	1 2 3
3 1 2	1 2 3	2 3 1

Square (1) comprises the elements in the first position, square two comprises the elements in the second position, and square three comprises the elements in the third position. The tables are constructed using a one-step cyclic permutation of the numbers in the design. That is, the next design in the sequence is constructed by moving the first number of the current design to the extreme right, simultaneously moving all other numbers to the left (Winer, 1962). A composite design constructed from cyclic elements is balanced in the sense that each attribute level appears once and only once in a cell and once and only once in each position in a row or column. Thus, a row or column of a composite design represents a balanced Pareto optimal subset; i.e., every attribute and attribute level is represented and no alternative dominates another.

The remaining alternatives of Table 1 can be constructed by interchanging the second and third arrays of Table 2 to yield a new composite design:

		COLUMNS		
ROWS	1	1 3 2	2 1 2	3 2 1
	2	2 1 3	3 2 1	1 3 2
	3	3 2 1	1 3 2	2 1 3

The first row of the above design represents alternatives 2, 3, and 6 respectively of Table 1. With the addition of the alternative comprised of the second level of each attribute, the set will remain Pareto optimal, but it will no longer be balanced. The set of seven alternatives will be exhaustive, however, in the sense that each of the remaining 20 alternatives that can be constructed from the 3³ design will be dominated, or will dominate at least one of the seven alternatives in the Pareto optimal set. In this sense the seven alternatives represent the largest Pareto optimal subset that can be constructed from a 3³ design.

Constructing Sets for Larger Designs

While Pareto optimal subsets can be constructed for larger designs following generalizations of the procedures discussed above, the exercise is cumbersome and can become quite laborious. However, if we note that each attribute level will occur once and only once in each alternative generated by such procedures, a simple, easily programmable rule can be explicated which facilitates identification of desired alternatives.

This is the rule. For p^p designs select alternatives whose attribute levels sum to $\binom{p+1}{2}$. Since each level occurs once in an alternative, the sum of the levels will be the same for all alternatives. For example, in Table 2, if the letter (A) is replaced with the value one, (B) with the value two, and (C) with the value three, it is clear that for each of the six alternatives in Table 1, the sum of the attribute levels is six. Likewise, the desired alternatives from a 4⁴ design will have attribute levels which sum to ten, the desired alternatives from a 5⁵ design will have attribute levels which sum to 15, and so forth. The Pareto optimal alternatives based on cyclic hyper-graeco-latin squares will form a subset of the alternatives generated following the above rule. The remaining alternatives

identified are those in which a given attribute level appears more than once. For example, the alternative characterized by the second level of each attribute (2,2,2) of a 3^3 design is such an alternative. The second level appears three times, but the sum of these levels is equal to six, as required by the rule.

Table 3 presents Pareto optimal subsets for the 3^3 , 4^4 , and 5^5 factorial series. The sets are complete for the 3^3 and 4^4 series. As indicated in the above discussion, the 3^3 series contains seven alternatives, six of which are balanced in the sense that each level appears once and only once in an alternative. The 4^4 series contains 44 alternatives, 24 of which are balanced. The 5^5 series contains 381 elements. Only the 120 balanced elements are reported in order to conserve space. The 6^6 series is presented in Table 4. There are 4332 alternatives in this series, 720 of which are balanced. Only the seed balanced alternatives are listed.

Organizing According to BIBD's

Even restricted to Pareto optimal subsets, collecting choice data vis-a-vis the selected alternatives can become a formidable task. Twenty-one pairwise comparisons are required for the 7 alternatives in the 3^3 series. Two hundred and seventy-six pairwise presentations would be required for the 24 balanced alternatives in the 4^4 series. However, the 24 alternatives can be conveniently organized into six blocks of four alternatives each, such that the incidence of attribute levels is balanced within blocks and within the entire design. The following blocks illustrate this property.

Alternative	Block					
	(1)	(2)	(3)	(4)	(5)	(6)
1.	1234	1243	1324	1342	1423	1432
2.	2341	2431	3241	3421	4231	4321
3.	3412	4312	2413	4213	2314	3214
4.	4123	3124	4132	2134	3142	2143

Note that the blocks are constructed by cyclic permutation using the first six balanced elements of the 4^4 series (listed in Table 3). The balanced alternatives in the 5^5 series can be organized into 24 blocks of size five following similar procedures. The 6^6 series required 120 blocks of size six. Only a small fraction should be assigned a single subject and these should be replicated. Most subjects will find ranking six Pareto optimal 6 attribute alternatives a challenge.

In each of the above cases, the subject is required to rank the alternatives within blocks in order of preference. The resulting data can be analyzed "non-metrically," or following procedures suggested by Gulliksen and Tucker (1961), McKeon (1961), Wiley, MacLachlan, and Moynour (1976), or Green and Carmone (1977).

Table 3

BALANCED INCOMPLETE BLOCK DESIGNS

3^3 Design			4^4 Design			
1	2	3	1	1	4	4
1	3	2	1	2	3	4
2	1	3	1	2	4	3
2	2	2	1	3	2	4
2	3	1	1	3	3	3
3	1	2	1	3	4	2
3	2	1	1	4	1	4
			1	4	2	3
			1	4	3	2
			1	4	4	1
			2	1	3	4
			2	1	4	3
			2	2	2	4
			2	2	3	3
			2	2	4	2
			2	3	1	4
			2	3	2	3
			2	3	3	2
			2	3	4	1
			2	4	1	3
			2	4	2	2
			2	4	3	1
			3	1	2	4
			3	1	3	3
			3	1	4	2
			3	2	1	4
			3	2	2	3
			3	2	3	2
			3	2	4	1
			3	3	1	3
			3	3	2	2
			3	3	3	1
			3	4	1	2
			3	4	2	1
			4	1	1	4
			4	1	2	3
			4	1	3	2
			4	1	4	1
			4	2	1	3
			4	2	2	2
			4	2	3	1
			4	3	1	2
			4	3	2	1
			4	4	1	1

5^5 Design

1	2	3	4	5	2	4	5	1	3	4	2	3	1	5
1	2	3	5	4	2	4	5	3	1	4	2	3	5	1
1	2	4	3	5	2	5	1	3	4	4	2	5	1	3
1	2	4	5	3	2	5	1	4	3	4	2	5	3	1
1	2	5	3	4	2	5	3	1	4	4	3	1	2	5
1	2	5	4	3	2	5	3	4	1	4	3	1	5	2
1	3	2	4	5	2	5	4	1	3	4	3	2	1	5
1	3	2	5	4	2	5	4	3	1	4	3	2	5	1
1	3	4	2	5	3	1	2	4	5	4	3	5	1	2
1	3	4	5	2	3	1	2	5	4	4	3	5	2	1
1	3	5	2	4	3	1	4	2	5	4	5	1	2	3
1	3	5	4	2	3	1	4	5	2	4	5	1	3	2
1	4	2	3	5	3	1	5	2	4	4	5	2	1	3
1	4	2	5	3	3	1	5	4	2	4	5	2	3	1
1	4	3	2	5	3	2	1	4	5	4	5	3	1	2
1	4	3	5	2	3	2	1	5	4	4	5	3	2	1
1	4	5	2	3	3	2	4	1	5	5	1	2	3	4
1	4	5	3	2	3	2	4	5	1	5	1	2	4	3
1	4	5	3	4	3	2	5	1	4	5	1	3	2	4
1	5	2	4	3	3	2	5	4	1	5	1	3	4	2
1	5	3	2	4	3	4	1	2	5	5	1	4	2	3
1	5	3	4	2	3	4	1	5	2	5	1	4	3	2
1	5	4	2	3	3	4	2	1	5	5	2	1	3	4
1	5	4	3	2	3	4	2	5	1	5	2	1	4	3

5⁵ Design (continued)

2 1 3 4 5	3 4 5 1 2	5 2 3 1 4
2 1 3 5 4	3 4 5 2 1	5 2 3 4 1
2 1 4 3 5	3 5 1 2 4	5 2 4 1 3
2 1 4 5 3	3 5 1 4 2	5 2 4 3 1
2 1 5 3 4	3 5 2 1 4	5 3 1 2 4
2 1 5 4 3	3 5 2 4 1	5 3 1 4 2
2 3 1 4 5	3 5 4 1 2	5 3 2 1 4
2 3 1 5 4	3 5 4 2 1	5 3 2 4 1
2 3 4 1 5	4 1 2 3 5	5 3 4 1 2
2 3 4 5 1	4 1 2 5 3	5 3 4 2 1
2 3 5 1 4	4 1 3 2 5	5 4 1 2 3
2 3 5 4 1	4 1 3 5 2	5 4 1 3 2
2 4 1 3 5	4 1 5 3 2	5 4 2 1 3
2 4 1 5 3	4 1 5 3 2	5 4 2 3 1
2 4 3 1 5	4 2 1 3 5	5 4 3 1 2
2 4 3 5 1	4 2 1 5 3	5 4 3 2 1

Table 4
SEED VALUES 6⁶ DESIGN

1 2 3 4 5 6	1 3 5 6 2 4	1 5 3 4 2 6
1 2 3 4 6 5	1 3 5 6 4 2	1 5 3 4 6 2
1 2 3 5 4 6	1 3 6 2 4 5	1 5 3 6 2 4
1 2 3 5 6 4	1 3 6 2 5 4	1 5 3 6 4 2
1 2 3 6 4 5	1 3 6 4 2 5	1 5 4 2 3 6
1 2 3 6 5 4	1 3 6 4 5 2	1 5 4 2 6 3
1 2 4 3 5 6	1 3 6 5 2 4	1 5 4 3 2 6
1 2 4 3 6 5	1 3 6 5 4 7	1 5 4 3 6 2
1 2 4 5 3 6	1 4 2 3 5 6	1 5 4 6 2 3
1 2 4 5 6 3	1 4 2 3 6 5	1 5 4 6 3 2
1 2 4 6 3 5	1 4 2 5 3 6	1 5 6 2 3 4
1 2 4 6 5 3	1 4 2 5 6 3	1 5 6 2 4 3
1 2 5 3 4 6	1 4 2 6 3 5	1 5 6 3 2 4
1 2 5 3 6 4	1 4 2 6 5 3	1 5 6 3 4 2
1 2 5 4 3 6	1 4 3 2 5 6	1 5 6 4 2 3
1 2 5 4 6 3	1 4 3 2 6 5	1 5 6 4 3 2
1 2 5 6 3 4	1 4 3 5 2 6	1 6 2 3 4 5
1 2 5 6 4 3	1 4 3 5 6 2	1 6 2 3 5 4
1 2 6 3 4 5	1 4 3 6 2 5	1 6 2 4 3 5
1 2 6 3 5 4	1 4 3 6 5 2	1 6 2 4 5 3
1 2 6 4 3 5	1 4 5 2 3 6	1 6 2 5 3 4
1 2 6 4 5 3	1 4 5 2 6 3	1 6 2 5 4 3
1 2 6 5 3 4	1 4 5 3 2 6	1 6 3 2 4 5
1 2 6 5 4 3	1 4 5 3 6 2	1 6 3 2 5 4
1 3 2 4 5 6	1 4 5 6 2 3	1 6 3 4 2 5
1 3 2 4 6 5	1 4 5 6 3 2	1 6 3 4 5 2
1 3 2 5 4 6	1 4 6 2 3 5	1 6 3 5 2 4
1 3 2 5 6 4	1 4 6 2 5 3	1 6 3 5 4 2
1 3 2 6 4 5	1 4 6 3 2 5	1 6 4 2 3 5
1 3 2 6 5 4	1 4 6 3 5 2	1 6 4 2 5 3
1 3 4 2 5 6	1 4 6 5 2 3	1 6 4 3 2 5
1 3 4 2 6 5	1 4 6 5 3 2	1 6 4 3 5 2
1 3 4 5 2 6	1 5 2 3 4 6	1 6 4 5 2 3
1 3 4 5 6 2	1 5 2 3 6 4	1 6 4 5 3 2
1 3 4 6 2 5	1 5 2 4 3 6	1 6 5 2 3 4
1 3 4 6 5 2	1 5 2 4 6 3	1 6 5 2 4 3
1 3 5 2 4 6	1 5 2 6 3 4	1 6 5 3 2 4
1 3 5 2 6 4	1 5 2 6 4 3	1 6 5 3 4 2
1 3 5 4 2 6	1 5 3 2 4 6	1 6 5 4 2 3
1 3 5 4 6 2	1 5 3 2 6 4	1 6 5 4 3 2

References

Anderson, N. H., Information Integration Theory: A Brief Survey, Center for Human Information Processing, University of California, La Jolla, California, April 1972.

Bettman, J. R., N. Capon, and R. J. Lutz, "Cognitive Algebra in Multiattribute Models," Journal of Marketing Research, 12 (May 1975), pp. 151-64.

_____, "Multiattribute Measurement Models and Multiattribute Theory: A Test of Construct Validity," Journal of Consumer Research (March 1975), pp. 1-85.

Dawes, R. M., and B. Corrigan, "Linear Models in Decision Making," Psychological Bulletin, 1974, pp. 95-106.

Fiedler, J. A., "Optimizing Product Design Through Analysis of Consumer Trade Off's," M. Venkateson (ed.), Proceedings 3rd Annual Conference Association for Consumer Research, 1972, pp. 279-93.

Green, P. E., "On the Design of Choice Experiments Involving Large Numbers of Factors or Factor Levels." In S. Ward and P. Wright (eds.), Proceedings of the 4th Annual Conference of the Association for Consumer Research, 1974, pp. 228-41.

_____, and F. J. Carmone, "A BIBD/Logit Approach to Conjoint Measurement," University of Pennsylvania Working Paper, April 1977.

Green, P. E., F. J. Carmone, and Y. Wind, "Subjective Evaluation Models and Conjoint Measurement," Behavioral Science, Vol. 17 (May 1972), pp. 288-99.

Green, P. E. and V. R. Rao, "Conjoint Measurement Method of Quantifying Judgmental Data," Journal of Marketing Research, Vol. III (August 1971), pp. 355-62.

Green, P. E., and Y. Wind, Multiattribute Decisions in Marketing: A Measurement Approach. Hinsdale, Illinois: The Dryden Press, 1973.

Gulliksen, H. and L. R. Tucker, "A General Procedure for Obtaining Paired Comparisons from Multiple Rank Orders," Psychometrika, 26 (June 1961), pp. 173-83.

Johnson, R. M., "Trade-off Analysis of Consumer Values," Journal of Marketing Research, Vol. XI (May 1974), pp. 121-27.

Krantz, D. and A. Tversky, "Conjoint-measurement Analysis of Composition Rules in Psychology," Psychological Review (March 1971), pp. 151-69.

Kruskal, J. R., "Analysis of Factorial Experiments by Estimating Monotone Transformations of the Data," Journal of the Royal Statistical Society, B, 27, 1965, pp. 251-63.

McKeon, J. J., 1961. "Measurement Procedures Based on Comparative Judgement," unpublished doctoral dissertation, University of North Carolina, Chapel Hill.

Wiley, J. B., D. L. MacLachlan, R. Moinpour, "Comparison of Stated and Inferred Parameter Values in Additive Models: An Illustration of a Paradigm," William Perrault, Jr. (ed), Advances in Consumer Research, Vol. IV, pp. 98-105.

Winer, B. J., Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.

FOUR METHODOLOGICAL PROBLEMS IN MULTI-ATTRIBUTE
ATTITUDE MODELS¹

Michel Laroche, Simmons College

Abstract

Four methodological problems are identified: two dealing with the scaling decision, i.e., origin and range; and two dealing with the relation among attributes, i.e., interaction and halo effects. Two procedures are proposed in order to improve the specification of the multi-attribute attitude models and our understanding of the process of attitude formation, as well as their predictive abilities.

Identification of the Problems

There has been a great amount of research in marketing on multi-attribute attitude models inspired by the work in social psychology of Rosenberg (1956) and Fishbein (1967). Basically, consumers evaluate competing brands on a certain number n of attributes according to the following multiplicative rule:

$$A_j = \prod_{i=1}^n a_i B_{ij} \quad (1)$$

where:

A_j is the consumer's attitude toward brand j ;

B_{ij} is the strength of belief i about brand j , i.e., the probability that brand j is associated with attribute i ;

a_i is the evaluation of attribute i .

Some problems appeared, as researchers tried to implement the previous model or to compare its performance with additive or other nonlinear models. Excellent reviews of these attempts are available in Wilkie and Pessemier (1973) and in Einhorn and Gonedes (1971). The four major problems could be described as follows:

1. The problem of the origin of the scales used in measuring B_{ij} and a_i . As it was correctly pointed out by Schmidt and Wilson (1975), by choosing between unipolar (for example, from 1 to 5) and bipolar (from -2 to +2) scales for each concept, the researcher is able to manipulate the correlation between A_j and the summation of $a_i B_{ij}$'s. This is so because (1) is basically a nonlinear model, and a transformation of its variables by manipulation of the scales is likely to modify the functional form by which it is represented. If we note that the coefficient of correlation between two entities is unaffected by linear transformation of these entities, we can see that this problem of the origin of the scales will not affect the performance of linear models of attitude formation.

2. The problem of the range of the scales used in measuring B_{ij} and a_i . Most of the published studies have used the same range for all the concepts, usually scales with 5 or 7 points. As Etter (1975) correctly remarked, the model (1) will be sensitive to variations in the range of these scales, and the introduction of scale transformation parameters could improve its per-

formance. This is so for the same reason as before, i.e., (1) is basically a nonlinear model.

3. The model (1) does not take into account the interaction among the different attributes. In effect, the algebraic rule contained in (1) is equivalent to assuming independence of the different effects. It is evident that this is very often not the case. Sheth (1974) proposed a two-step procedure: a reduction of the dimensions by using a decomposition procedure of the different beliefs followed by a regression analysis in order to determine the weights of the underlying independent beliefs. But this method assumes homogeneity among the respondents and does not solve the problem of the origin. Besides, the derived dimensions are often difficult to interpret.

4. It has been observed that respondents tend to indicate higher beliefs about a brand according to the general attitude toward this brand. This phenomenon is the halo effect, and it represents a serious methodological problem in multi-attribute attitude models (see Beckwith and Lehmann, 1975).

This paper will attempt to provide some alternatives in order to obviate some of the problems which have just been identified.

The Problem of the Origin

When studying consumer behavior, the researcher selects a scale in order to measure a given phenomenon, such as the strength of a belief or the evaluation of an attribute. The phenomenon is assumed to exist independently of the type of scale used. For example, we feel the same way today, whether the temperature is expressed in Celsius or Fahrenheit. Faced with bipolar adjectives, the respondent indicates his position by checking one of the possible alternatives. Then, the researcher attributes numbers to these alternatives. As long as this measure is used in an additive way, no problem arises. But in multi-attribute attitude models, researchers are trying to multiply such measures and in order to do so ratio scales are needed. In physics, absolute temperatures had to be introduced when temperature was one of the multiplying variables in a given law. Similarly, it is necessary here to determine the position of the origin of the scale for each concept. It becomes a parameter to be estimated when applying the model (1), in the same way as regression parameters are obtained. That is, the origins for the a_i 's and the B_{ij} 's are the ones which maximize the correlation coefficient between the calculated and the measured attitudes.

This approach is superior to the arbitrary determination of these parameters which has been found in the literature. For example, completely different predictions could result from such an initial choice, as is illustrated in Table 1 with an hypothetical example. As is evident in this case, the determination of the origin is not only a question of calibrating a tool; it is essentially a very important element of model specification.

¹This research is part of a project supported by the Laboratoire de Recherche of the Faculté des Sciences de l'Administration of the Université Laval, Québec.

TABLE 1

With the same data, different predictions are made according to the scales chosen.

		i = 1	i = 2	A _j
Bipolar model	B _{i1}	0	1	1
	B _{i2}	-1	0	2*
	a _i	-2	1	
Mixed models	B _{i1}	3	4	-2
	B _{i2}	2	3	-1*
	a _i	-2	1	
	B _{i1}	0	1	4*
	B _{i2}	-1	0	-1
	a _i	1	4	
Unipolar model	B _{i1}	3	4	19*
	B _{i2}	2	3	14
	a _i	1	4	

* Highest predicted attitude

In order to illustrate one of the procedures proposed in this paper, a set of data was collected for 75 persons selected in a shopping center of Québec. The questionnaire used dealt with the choice between two competing brands of cola in the large size category: the 53 oz. bottle of Pepsi and the 40 oz. bottle of Coke. Four criteria were used: economy, ease of manipulation, brand reputation and ease of storage in the refrigerator. The bipolar adjectives used were "bad description good/description" for the strengths of beliefs and "not important at all/very important" for the evaluations. Both were collected on a 5-points scale. The raw data were then transformed by varying the origin for each set of scales, and the transformed data were used in (1) in order to obtain the predicted attitude which was correlated with the measured overall attitude. The correlation coefficients thus obtained are presented in Table 2. The results illustrate dramatically the effect of changing the codifications of the data used in (1).

In effet, the model (1) should be rewritten as

$$A_j = \sum_{i=1}^n (a_i - a)(B_{ij} - B)$$

where a and B are constants to be determined.

In Table 2, one can find the correlation for the bipolar model (a = 3, B = 3), the mixed models (a = 0, B = 3; a = 3, B = 0) and the unipolar model (a = 0, B = 0). Further analysis leads to the following conclusions:

1. The best model for Pepsi alone is a = 0, B = 4.4 for which the correlation coefficient is equal to .6450;

2. The best model for Coke alone is a = 2.1, B = 3.5 for which the correlation coefficient is equal to .6188;

3. The best model for both is a = 1.1, B = 3.8 for which the coefficients of correlation are respectively .6277 and .5942. This was obtained by maximizing the sum of both correlation coefficients.

Finally, it can be shown that Anderson's averaging model is relatively unaffected by the problem of the origin. This would explain why this model performs in a more consistent manner than the model (1). Basically, one can generally describe the averaging model as follows (Anderson, 1970):

The evaluation scale is a weight variable w_i, which one can define as

$$w_i = \frac{a_i - a}{\sum_{i=1}^n (a_i - a)}, \quad (3)$$

which leads to

$$\sum_{i=1}^n w_i = 1.$$

These weights then combine in a multiplicative fashion with the strength of belief in order to yield the attitude A'_j:

$$A'_j = \sum_{i=1}^n w_i (B_{ij} - B) \quad (4)$$

By decomposing (4) and rearranging terms, one finds:

$$A'_j = \sum_{i=1}^n w_i B_{ij} - B.$$

This last expression shows that the coefficient of correlation between A'_j and the measured attitude would be independent of B, which is the origin of the belief scale. But the correlation coefficient is still dependent upon a, which is the origin of the evaluation scale. Intuitively, one would expect a to be the lowest value of that scale. This was borne out experimentally on the same set of data as before. The procedure was the same as before: a was varied, and for each value of a the correlation coefficient between predicted and measured attitudes was computed. The origin a was determined to be equal to 1, and the corresponding coefficients of correlation were respectively for Pepsi and for Coke equal to .6436 and .6114. It is worth noting here that these coefficients are very close to the ones obtained previously for Fishbein's model. This might suggest that, with proper adjustment of the scales, their predictive abilities are similar.

Limitation of the method

This procedure for determining the origin of each type of scale, i.e., a_i and B_{ij}, suffers from one implicit assumption: that all attributes have the same origin on the same scale. Although this is intuitively appealing, it should be tested, and the previous method could be adapted in order to fit the following model:

$$A''_j = \sum_{i=1}^n (a_i - a_{oi}) (B_{ij} - B_{oi}). \quad (5)$$

This would necessitate the estimation of 2n parameters instead of 2.

TABLE 2

Variations of the coefficient of correlation between calculated and measured attitudes as a function of the origins of the evaluation and belief scales for Pepsi and Coke.

	a = 0	a = 1	a = 2	a = 3	a = 4	a = 5
B = 0	0.4141 0.4371	0.3250 0.4048	0.1674 0.3409	-0.0727 0.2281	-0.3153 0.0740	-0.4670 -0.0742
B = 1	0.4931 0.4837	0.4275 0.4676	0.2819 0.4181	-0.0114 0.2937	-0.3292 0.0906	-0.4888 -0.0938
B = 2	0.5661 0.5269	0.5324 0.5334	0.4300 0.5167	0.1035 0.3960	-0.3157 0.1130	-0.4818 -0.1156
B = 3	0.6193 0.6567	0.6081 0.5826	0.5466 0.6047	0.2450 0.5088	-0.2294 0.1286	-0.4242 -0.1304
B = 4	0.6435 0.5626	0.6307 0.5879	0.5603 0.6023	0.2945 0.4778	-0.1047 0.1111	-0.3277 -0.1287
B = 5	0.6402 0.5415	0.6099 0.5440	0.5148 0.5076	0.2878 0.3473	-0.0131 0.0794	-0.2288 -0.1151

The Problem of the Range of the Scales

The same arguments as for the origin can be used to defend the need for a determination, at least implicitly, of the range of the scales used to measure a_i and B_{ij} . But there is one major exception for which this problem does not arise. If one assumes that all the a_i scales have the same range r_1 , and that all the B_{ij} scales have the same range r_2 , then the analysis of the behavior of the correlation coefficient shows that any uniform change of range for each group would not affect it. This result is a major difference with the previous problem.

On the other hand, if the ranges of, say, the a_i scales depend on the nature of the attribute, they become parameters to be estimated. This is what was done implicitly in the second phase of Sheth's method, and his regression coefficients represent the product of the average evaluation by a range adjustment factor (Sheth, 1974). One can say that, in general, one of the main advantages of multiple regression in this context is its ability to adjust for the misspecification of the range of each scale. This property will be used later in conjunction with the method of orthogonal polynomials.

Interactions Among Attributes and the Halo Effect

The model represented in (1) assumes that each attribute is processed by the consumer independently of the other ones. This would justify the manipulation of a selected attribute in order to improve the attitude toward a given brand. However, this is not often the case. For most products it is likely to find high correlations among attributes on the a_i and B_{ij} dimensions. For example, in the case of these soft drinks, we should expect an interaction between the manipulation and the storage criteria. Then, the question in terms of attitude change becomes one of direction and degree of modification of the related components of (1). If an induced variation in one of the attributes produces changes in the same direction on all the related attributes, then there will be some positive attitude change. But, if the induced variation produces opposite changes, then it will be necessary to determine, if possible and if desirable, the approximate amount of stimulation that would lead to positive attitude change.

A particular case of interaction effects among attributes is due to the halo effect (Beckwith and Lehmann, 1975). In their paper, Beckwith and Lehmann suggest the introduction in the application of the model (1) of the average attitude A^* as a proxy variable for this effect. It can be argued that a better proxy variable for the halo effect is the attribute of brand reputation. The higher the reputation of the brand, the more likely it will be overvalued on certain attributes. The main advantage of using this variable is that in doing so we better capture individual differences. In addition, it is possible to study the interactions involving the halo effect as will be shown in the next section.

Application of the Method of Orthogonal Polynomials

Methodology

The method of orthogonal polynomials developed by Laroche (1974) from the work of Fisher (1958) is particularly useful in the context of empirically determining the consumer's decision rules. Its major advantages are as follows:

1. It adapts itself to the two problems associated with the scales, i.e., origin and range.
2. It decomposes the effects into independent entities: main and interaction.
3. All interactions can be introduced into the analysis.
4. The halo effect can be determined more precisely.

The same data as before are used in this application. For each attribute, the main effects are calculated (linear, quadratic and cubic). To summarize, the method consists of decomposing the effect of x on y into a trend or linear component, a quadratic component uncorrelated with the trend, and a cubic component uncorrelated with the previous two. On the interaction side, and this is the important part, the effects of x and z on y are decomposed into two main effects or trends and an interaction component uncorrelated with the trends of x and z . Thus, for each set of attributes, an interaction component is calculated. All of

TABLE 3

Regression Results for Pepsi and Coke Using the Method
of Orthogonal Polynomials

Brand	Regression Coefficients (t-values)								Constant	R
	(1) Economy		(2) Manipulation		(3) Brand Reputation		(4) Storage			
Pepsi	P_1	$P_1 \times a_1$	P_3	$P_3 \times a_2$	a_3	$a_3(\text{cubic})$	$P_3 \times a_4$	$a_3 \times a_4$	(e)	(F)
	.325	.862	.711	2.78	-.328	-10.91	-3.21	-1.93	.6026	.8118
	(2.73)	(1.76)	(7.90)	(3.19)	(-2.65)	(-3.38)	(-4.11)	(-3.83)	(.1909)	(16.2)
Coke			C_3	$C_3 \times a_2$			$C_4 \times a_4$			
			.740	-.996			.815		.6974	.7160
			(7.71)	(-2.32)			(2.52)		(.2022)	(25.2)

where a_1, a_2, a_3, a_4 are the evaluations; P_1, P_2, P_3, P_4 are the beliefs for Pepsi;

C_1, C_2, C_3, C_4 are the beliefs for Coke.

these components are then entered into a stepwise regression, and the resulting coefficients are presented in Table 3.

Major findings

1. The criterion of economy is independent of the other criteria, including the halo effect. It is significant only for the most economical brand, i.e., the 53 oz. bottle of Pepsi. On that dimension, since the only significant terms are the strength of belief (main) and the interaction (bipolar), the model used by the consumers is a mixed one: roughly bipolar on the evaluation (a_1) and roughly unipolar on the strength of belief (P_1). This is consistent with the previous overall result.

2. The previous finding for a_1 is also true for a_2 and a_4 for both brands.

3. For both brands, the reputation of each one on the main effect is the most important criterion. The halo effect is apparent by the significant interaction of brand reputation with manipulation in the case of Coke and with manipulation and storage in the case of Pepsi. In particular, for Coke the more important the criterion of manipulation is, the lower the attitude toward Coke is. Similarly for Pepsi, the more important to the consumer both attributes of manipulation and storage are, the lower the attitude toward Pepsi will be.

4. Among all the nonlinear components introduced in the stepwise regression, the only significant one is a_3 (cubic). This last term, coupled with the linear trend a_3 , shows that there is a complex negative reaction against the brand Pepsi since the more important this criterion is, the lower the attitude toward Pepsi is.

5. In the case of Coke, the criterion of storage seems to follow a bipolar model. In fact, for both products and for all the attributes, this is the only case for which the decision rule is close to the one suggested by Fishbein (1967).

Discussion

This last empirical approach to the determination of the process of attitude formation is probably more fruitful than the comparison of several possible models which could be premature. A strategy of accumulating empirical results of the kind presented in this paper would provide researchers the raw material upon which to build better models of attitude formation or to more adequately adapt existing models or formalize existing theories. This task would also require a more explicit elicitation of the basic assumptions upon which models are elaborated. A case in point is the assumptions underlying the scaling decision, the multiplicative rule, and the additivity of the products.

Summary

Four methodological problems were identified in the context of multi-attribute attitude models, more precisely the linear compensatory model. Two were related to the scaling decision, i.e., the origin and the range of the different scales used to measure the evaluations and the strengths of beliefs. A procedure was proposed in order to obviate these problems. It was based upon a numerical search with the correlation coefficient as the objective function. Two other problems were related to the relationship among attributes, i.e., the interaction and the halo effects. It was proposed that the application of the method of orthogonal polynomials would be very useful in order to incorporate and analyze all four issues. The empirical results are very promising, and they suggest that a better understanding of attitude formation would be reached if we accumulated such findings. These would provide the building blocks for constructing better models, both in terms of explanation and prediction.

References

- Anderson, N. H. "Integration Theory and Attitude Change," *Psychological Review*, 77 (May 1970), 153-170.
- Beckwith, N. E. and D. R. Lehmann, "The Importance of Differential Weight in Multiple Attribute Models of Consumer Attitude", *Journal of Marketing Research*, Vol. X (May 1973), 141-145.

Beckwith, N. E. and D. R. Lehmann, "The Importance of Halo Effects in Multi-Attribute Attitude Models", Journal of Marketing Research, Vol. XII (August 1975), 265-275.

Einhorn, H. J. and N. J. Gonedes, "An Exponential Discrepancy Model for Attitude Evaluation", Behavioral Science, 16 (March 1971), 152-157.

Etter, W. L., "Attitude Theory and Decision Theory: Where is the Common Ground?" Journal of Marketing Research, Vol. XII (November 1975), 481-483.

Fishbein, M., "A Behavior Theory Approach to the Relations Between Beliefs About an Object and the Attitude Toward the Object", in M. Fishbein, ed., Readings in Attitude Theory and Measurement (New York: Wiley, 1967), 389-400.

Fisher, R. A., Statistical Methods for Research Workers, 13th ed. (New York: Hefner, 1958).

Laroche, M., "A New Approach to Nonlinear Consumer Behavior and Market Segmentation by the Use of Orthogonal Polynomials", Unpublished Ph.D. Dissertation, Columbia Graduate School of Business, 1974.

Rosenberg, M. J., "Cognitive Structure and Attitudinal Affect", Journal of Abnormal and Social Psychology, 53 (November 1956), 367-372.

Sheth, J. N., "A Field Study of Attitude Structure and Attitude-Behavior Relationship", in J. N. Sheth, ed., Models of Buyer Behavior (New York: Harper & Row, 1974).

Wilkie, W. L. and E. A. Pessemier, "Issues in Marketing's Use of Multi-Attribute Attitude Models", Journal of Marketing Research, 10 (November 1973), 428-441.

MEASURED ATTRIBUTE WEIGHTS CAN MAKE A DIFFERENCE

Richard A. Werbel,
University of Illinois at Circle Campus

Abstract

This study compares the correlation between stated attitudes and attitudes generated by the linear compensatory model using measured attribute weights to the same correlation using uniform weights. Unlike previous studies, attempts were made to minimize both an implicit weighting effect and a positive correlation between or among beliefs. In addition, a new approach of measuring attribute importance was used to minimize any systematic response bias. The correlations were substantially greater with measured weights than uniform weights.

Introduction

A large number of studies have been conducted which compare a compensatory multi-attribute model with uniform importance weights to one with measured weights.¹ Bass and Wilkie (1973) summarized the results of 15 earlier studies and other studies have been published since this review (for example, Dubois, 1974; Holbrook and Hulbert, 1974; Houston and Gillespie, 1974; Pekelman and Sen, 1974; Wildt and Bruno, 1974; and Mazis, Ahtola, and Klippel, 1975). This author was able to find only one study in which the correlation between stated and generated attitudes was dramatically higher with measured weights (Pekelman and Sen, 1974). Thus, the evidence to date indicates that uniform weights can predict attitudes about as well as measured weights.

Yet, it might be premature to conclude that uniform weights actually are used by consumers in forming attitudes. First, and probably most important, attitudes generated by a uniform weights model may be quite similar to those generated by measured weights, even when there is a large difference between measured and uniform weights. Second, methods used to measure weights may suppress true differences between measured and uniform weights. Both of these conditions are discussed in the following section.

In this study, attempts were made to increase the probability that attitudes generated by uniform weights differed from those generated by measured weights when a large difference between uniform and measured weights existed. In addition, a new procedure for measuring attribute importance was used. It was expected that this method might generate more accurate weights than methods used previously.

It has been argued that there is little reason to be concerned with attribute importance because the "value" component of Fishbein's model incorporates both attribute importance and the degree of deviation from an ideal point (Cohen, Fishbein, and Ahtola, 1972). Although it is likely that attribute importance is incorporated into this component, it seems desirable from a diagnostic standpoint to attempt to isolate all attitude components. For example, it might be relevant to know the distance between the ideal point and a cognitive belief to determine the likelihood of changing the cognitive belief.

Explanations for the Inability Of Measured Weights to Outperform Uniform Weights

If a belief toward a brand on one attribute is reasonably similar in an affective and relative sense to beliefs toward the same brand on other salient attributes, a model with weights substantially different from uniformity will tend to generate the same attitudes as a model with uniform weights (Beckwith and Lehmann, 1973). Although it is difficult to identify the number of studies in which this problem has existed because the correlation usually has not been reported, only Holbrook and Hulbert (1974), and Pekelman and Sen (1974) seem to have in large eliminated it.

Sheth and Talarzyk (1972) argued that there may be a greater spread of beliefs toward various brands on an attribute that is important than on one that is unimportant. In this situation, since attitudes would tend to be consistent with affective beliefs on the attribute or attributes with the greatest spread when uniform weights are used, attitudes generated by uniform weights would tend to be similar to those generated by measured weights. This "implicit weighting" effect has been reported in previous studies (Sheth and Talarzyk, 1972; Beckwith and Lehmann, 1973; Wilkie and Weinrich, 1973; Holbrook and Hulbert, 1974). Perhaps it is most significant that Wilkie and Weinrich discovered an implicit weighting effect because cognitive beliefs were used rather than affective ones and because some of the attributes included were reasonably objective. Based upon these results, the implicit weighting effect might have existed in other studies even though it was not reported.

There is a second explanation for implicit weighting that has not been advanced previously. The use of a number of reasonably specific attributes underlying the basic attribute of quality along with a single price attribute results in greater weight being attached to quality when uniform weights are used. If some consumers combine perceptions on reasonably specific attributes into a single "quality" perception and if quality has greater weight than price, an implicit weighting effect will exist.

Beckwith and Lehmann (1973) indicated that measures of attribute weights might not coincide with true weights. Using regression estimated weights to generate attitudes, they examined the correlation between generated attitudes and the group of stated attitudes which were used to determine the regression weights. Since the difference between the average correlation coefficient using regression estimated weights and that using stated weights was not that large (.095) it can be assumed that stated weights were similar to regression estimated weights and/or the existence of such factors as a positive correlation among beliefs resulted in suppressing the influence of weights upon generated attitudes. Although Beckwith and Lehmann did not indicate the association between regression estimated weights and stated weights, the existence of both a positive correlation among beliefs and an implicit weighting effect indicates that there may have been a reasonably substantial meas-

urement error.

Scott and Wright (1976) and Hughes and Guerrero (1971) compared regression estimated weights with stated weights. In both of these studies, the difference between the two sets of weights was rather large. For example, Hughes and Guerrero found that the computed weight for the performance of automobiles was 57.3 compared with a stated weight of 43.6 using a base of 100. Furthermore, regression estimated weights were more extreme than stated weights in both studies. Thus, there is some evidence that the use of direct questions to measure attribute importance results in a systematic response error which reduces the difference between attitudes generated by uniform weights and those generated by true weights.

Methodology

Undergraduate students in two sections of Principles of Marketing at a large midwestern university were questioned about their attitudes, beliefs, and the importance attached to various evaluative criteria. The product category examined was automobiles. The specific cars included were the 1) Chevrolet Chevette, 2) Ford Pinto, 3) Chevrolet Monza, 4) Ford LTD, and 5) Chevrolet Malibu. Due primarily to one of the methods used to measure attribute importance described later in the paper, an interactive computer program was used for the questioning process.

The two attributes used in the study were quality and price. The selection of quality and price as attributes should be explained, as quality has not been used in any previous studies, and because usually more than two attributes have been included in previous studies.

Before selecting quality and price, the decision was made to only include two attributes. This was done for three reasons. First, the use of two attributes lessens the likelihood of a positive correlation among beliefs existing, as it is difficult to select negatively correlated attributes when more than two attributes are used.² Second, any increase in the number of attributes might reduce the differences in the weights between attributes, particularly when a constant sum approach is used to measure weights. This argument is based on the assumption that it is reasonably unlikely that a weight of zero will be assigned to an attribute by a respondent. Third, for the reason mentioned in the previous section, the use of two attributes lessens the likelihood of an implicit weighting effect.

Once the decision was made to use two attributes, quality and prices were selected because it was expected that they would be negatively correlated. Furthermore, because only two attributes were used, it was important to use a rather broad attribute, such as quality, which would incorporate more specific attributes.

²It was thought that the positive correlation could be minimized by selecting attributes that were likely of being negatively related. An overestimation of the degree of negative correlation would have likely resulted in uncorrelated beliefs, while the attempt to select uncorrelated beliefs might have resulted in positively correlated beliefs. Furthermore, although the use of negatively correlated beliefs tended to result in similar attitudes being generated by two sets of weights when the rank orders of the attribute weights were equivalent, their use tended to result in attitudes generated by uniform weights different than those generated by nonuniform weights.

Prior to answering any questions, respondents were given information dealing with the price, seating capacity, gas mileage, safety upon collision, and seating comfort of each car. A photograph of each car also was presented. This was done to lessen the likelihood of the existence of a positive correlation among beliefs because the low-priced objects were rated "low" on the nonprice information. Furthermore, the presentation of the price information made it possible to use objective price beliefs.

With one exception, an attempt was made to present accurate information. The exception was that all cars were depicted as equal in terms of gas mileage. If the correct information regarding gas mileage had been presented and if respondents had integrated this information into the quality dimension, the positive correlation between beliefs might have increased. The possibility of this information being integrated into a price dimension would have made it difficult to use objective price beliefs. It was thought that a positive correlation might be less likely of existing with objective price beliefs than with normalized or nonnormalized price beliefs.

Respondents were instructed to assume that all the information presented was accurate, including that pertaining to gas mileage. They were also told that they did not have to use this information if they did not feel it was relevant, and that they could use information dealing with attributes not included if they felt it to be relevant.

After the above instructions were presented, respondents were questioned about their attitude toward the act of purchasing five different cars. First, they were asked to rank the objects. Using a four-point scale, they then were asked to indicate the extent that they were less likely to purchase the brand ranked second than the brand ranked first. This question was repeated for the third and second, fourth and third, and fifth and fourth brands. Answers to these questions were used to construct an interval-type attitude scale.

Ten-point scales were used to measure beliefs. This particular scale was selected to lessen the probability of an implicit weighting effect by restricting respondent choice to some extent. At the same time, it was thought that this scale provided sufficient choice to allow the perceived relative differences between objects to be reflected. The scale endpoints were labeled "medium to higher quality," "lower quality," "lower priced" and "medium to higher priced." The term "medium" was used to encourage use of extreme scale values.

In addition to perceived or nonnormalized beliefs, two other types of beliefs were used in the analysis. Normalized beliefs with both price and quality were used along with objective beliefs with price.

Both the reason and method of normalization used in this study differ from previous studies. In this study, normalized beliefs were calculated to equalize the quality spread between the lowest and highest ranked objects with the price spread for each respondent. This was done to lessen the possibility of an implicit weighting effect. Although this type of normalization might alter the perceptions of respondents, its use might be justified when an implicit weighting effect exists because of the differential weighting of attributes. In addition, the sensitivity of generated attitudes to changes in the relative spread of beliefs across attributes could be examined with this type of normalization.

This normalization was accomplished by giving the object that was perceived as highest in quality, (lowest in price), a value of 1 and the object that was perceived

as lowest in quality, (highest in price), a value of 10. The criterion used in altering the beliefs of the other objects was the preservation of the relative differences in nonnormalized beliefs between objects.

Objective beliefs were used with price to lessen the possibility of the existence of a positive correlation between beliefs. Objective price beliefs were determined by converting actual prices into a ten-point scale in which the object with the lowest price was given a value of 1 and the object with the highest price a value of 10. The actual prices presented to respondents were \$2900 for the Chevette, \$3100 for the Pinto, \$3900 for the Monza, \$4700 for the LTD, and \$4900 for the Malibu. Objective price beliefs were 1.0, 1.9, 5.5, 9.1 and 10.0. Another method that might have been used would have involved equating the spread of objective price beliefs with the spread with either nonnormalized quality or price beliefs, while preserving the relative price differences indicated by the actual prices. After examining the initial results, the decision was made to use the procedure indicated primarily because it allowed one to examine the impact that a price spread greater than a quality spread had upon the correlation between stated and generated attitudes with automobiles.

Two measures were used to measure attribute importance. One approach involved asking respondents directly to provide a numeric weight for both quality and price which indicated the relative importance of these two attributes. Respondents also were instructed with this approach that the sum of the weights for quality and price had to be equal to 1.0. With the exception of using a constant-sum scale, this approach is virtually identical to those used in previous studies. A new approach was used to measure attribute importance because, as indicated previously, the above approach may not be valid. This new approach involved asking either question (1) or (2).

- (1) How much less would (Y) have to cost than (X) for you to be just as likely to purchase (Y) as (X)?
- (2) How much more would (X) have to cost than (Y) for you to be just as likely to purchase (X) as (Y)?

X and Y were two of the five brands included in the study. Although the specific brands included could change from respondent to respondent and/or question (2) to question (1), X always was perceived as higher in quality than Y. Respondents were given nine specific prices to choose among in answering these questions. Answers were scaled for further analysis. The lowest price had a scale value equal to 2 and the largest price had a scale value equal to 10. Question (1) or (2) was repeated for each respondent with a different pair of brands. Mean weights were then calculated. These mean weights were used in generating attitudes using "trade-off" weights.

Equations (1) and (2) were used to algebraically determine the weights for quality and price.

$$(B_{qyk} - B_{qyk} + 1) W_{qk} + B_{pxk} W_{pk} =$$

$$(B_{qyk} - B_{qyk} + 1) W_{qk} + B_{pyk} W_{pk} \quad (1)$$

$$W_{qk} + W_{pk} = 1.0 \quad (2)$$

Such that:

W_{qk} = weight of quality (q) for consumer k;

W_{pk} = weight of price (p) for consumer k;

B_{qyk} = consumer k's nonnormalized quality belief toward the higher quality brand included in question (1) or (2);

B_{qyk} = consumer k's nonnormalized quality belief toward the lower quality brand included in question (1) or (2);

B_{pxk} = scale value for price given by consumer k in response to question (1) or (2); and

B_{pyk} = scale value for price of base brand in question (1) or (2). This value was equal to 1.0 for all respondents.

It should be noted that the difference between B_{qyk} and

B_{qyk} influences the possible range of values that W_{qk}

and W_{pk} can assume. For example, if the difference

between B_{qyk} and B_{qyk} is equal to 1, W_{qk} cannot assume

a value less than .5 or greater than .9. This potential

problem was solved by asking questions (3) and (4).

3. If you had to buy either (brand perceived as highest in quality) or (brand perceived as second highest in quality) and (latter brand) costs _____ less than (former brand), which brand would you purchase?
4. If you had to buy either (brand perceived as lowest in quality) or (brand perceived as highest in quality) and (former brand) costs _____ less than (latter brand), which brand would you purchase?

Answers to these two questions were used to determine whether question (1) or (2) would be asked, and to determine the proper difference between X and Y in question (1) or (2) in terms of nonnormalized quality beliefs.

With both questions, the price changed from respondent to respondent as the specific quality difference between the two brands in the questions changed. This was done to establish a consistent range for W_{qk} and W_{pk} .

With question (3), the price was selected to determine whether W_{qk} was greater or less than approximately .67.

With question (4), the price was selected to determine whether W_{qk} was greater or less than approximately .30.

Respondents who selected the higher quality brand in both question (3) and (4) were asked question (1) if a pair of brands existed with which $B_{qyk} - B_{qyk}$ was be-

tween .5 and 2.49. Respondents who selected the lower quality brand in question (3) and the higher quality brand in question (4) were asked question (1) if a pair of brands existed with which $B_{qyk} - B_{qyk}$ was between

2.5 and 5.49. Although these two groups of respondents both were asked question (1), the allowable range of the difference between the quality beliefs of the two brands included in the question was different. This was done because the first group indicated in their answers to question (3) and (4) that the weight for quality was greater than .67, while the second group indicated that the weight for quality was between .30 and .67. Respondents who selected the lower quality brand in both question (3) and (4) were asked question (2) if a pair of brands existed with which $B_{qyk} - B_{qyk}$ was greater

than 5.49. Respondents in all of the above groups were not asked either question (1) or (2) if two brands did not exist with which $B_{qyk} - B_{qyk}$ was in the proper

range. Respondents who selected the higher quality brand in question (3) and the lower quality brand in question (4) were eliminated from the final analysis because this answer pattern indicated that the respondent's weights were quite unstable.

It was thought that this approach, which is termed the "tradeoff approach," might lessen or eliminate the response error with the "direct approach" primarily because the tradeoff approach does not directly ask respondents to evaluate the importance of an attribute. To clarify this reason, it is thought that when making a purchase decision, consumers do not need to think about the importance of an attribute unless a tradeoff exists. If a tradeoff exists, consumers might attempt to resolve the specific tradeoff, rather than directly evaluating the importance of an attribute. For example, if Brand "A" costs 10¢ more than Brand "B" and is perceived as somewhat higher in quality, a consumer will try to determine whether the perceived quality difference in favor of Brand "A" is large enough to compensate for its higher price.

The tradeoff approach was used rather than regression estimated weights primarily because a unique set of weights cannot be calculated using multiple regression when beliefs are positively or negatively correlated. A unique set of weights will exist with the tradeoff approach when beliefs are correlated. In addition, weights can be calculated using only two objects with the tradeoff approach and it is not necessary to use a holdout group of objects for model testing. This last factor is an advantage because of the practical problems involved in using a holdout group of objects (Pekelman and Sen, 1974).

Results of the Analysis

The analysis primarily consisted of using equation (3) to generate attitudes for each object used in the study.

$$A_{jk} = \sum_{i=1}^m W_{ik} |B_{ijk} - I_{ik}| \quad (3)$$

such that:

A_{jk} = consumer k's attitude toward object j,

W_{ik} = importance weight given attribute i by consumer k,

B_{ijk} = consumer k's cognitive belief toward object j on attribute i, and

I_{ik} = consumer k's ideal point on attribute i.³

Four different combinations of beliefs were used: 1) nonnormalized quality and price beliefs; 2) normalized quality and objective price beliefs; 3) nonnormalized quality and objective price beliefs; and 4) normalized quality and price beliefs. With each of these four

³In defining and determining the ideal points for quality and price, the attempt was made to isolate the use of price by consumers to determine quality beliefs through such factors as status or worksmanship, from its use as a primary economic criterion. Thus, the ideal point was defined as the most preferred level of an attribute, all other things being equal. With this definition, the ideal point with both quality and price was equal to 0.

combinations, three sets of weights were used; 1) direct weights; 2) tradeoff weights; and 3) uniform weights. In each of these 12 instances, Pearson product-moment correlation coefficients were calculated for each individual respondent. Mean coefficients were calculated and selective comparisons were made.

Although affective beliefs could have been used rather than cognitive beliefs and ideal points, the use of cognitive beliefs allowed the use of objective price beliefs. Furthermore, it is more consistent to use cognitive beliefs with the method of normalization used than affective beliefs. As mentioned earlier, this method should be used if the difference in the relative spread of beliefs across attributes exists because of the relative importance of the two attributes. This condition is more likely of existing when cognitive beliefs are used.

The decision to use the city block version of the model rather than the Euclidean distance version, or some other distance formulation, was a rather arbitrary one. Only one distance formulation was used because it was thought that the difference between the average correlation of generated and stated attitudes using one version of the model and the same correlation using another version of the model would not be affected by the particular distance formulation used.

88 students completed the questioning process. 19 of the 88 students were eliminated from the analysis because attribute weights could not be determined using the tradeoff approach. More specifically, they were eliminated either because their answers to questions (3) and (4) indicated that they had highly unstable weights, or because they did not meet the criterion used to select the brands to be included in questions (1), (2), (3), or (4). These students had an average Pearson product-moment correlation coefficient with uniform weights quite similar to that for the 69 respondents included in the analysis, but had a coefficient with direct weights about .18 less than that for the 69 respondents.

The means of Pearson product-moment correlation coefficients for the twelve versions of the model are presented in Table 1. In summarizing the correlations,

Table 1

MEAN OF INDIVIDUAL CORRELATIONS BETWEEN STATED AND GENERATED ATTITUDES

Combination of Beliefs Used	Tradeoff Weights	Direct Weights	Uniform Weights
Nonnormalized Quality and Price Beliefs	.6881 ^a	.6648	.3807
Normalized Quality and Objective Price Beliefs	.7017	.6408	.2024
Nonnormalized Quality and Objective Price Beliefs	.6393	.4822	-.1095
Normalized Quality and Price Beliefs	.7080	.7024	.2759

^a N = 69 in all instances

when other factors are held constant the correlations were substantially greater with direct and tradeoff weights than with uniform weights. The differences ranged from .2841, which was the difference between uniform and direct weights using nonnormalized beliefs for both quality and price, to .7488, which was the difference between uniform and tradeoff weights using nonnormalized quality beliefs and objective price beliefs.

The use of various combinations of beliefs to lessen an implicit weighting effect had a rather substantial impact on the difference between uniform and measured weights (i.e., the combination of direct and tradeoff weights). The implicit weighting effect was greatest when nonnormalized beliefs were used with both quality and price. Here the average difference between uniform and measured weights was .2958. The implicit weighting effect was least when nonnormalized quality beliefs and objective price beliefs were used.⁴ Here the difference between uniform and measured weights was .6703. The use of either normalized quality and price beliefs, or normalized quality and objective price beliefs resulted in an implicit weighting effect between the two extremes above. Here the difference between uniform and measured weights was .4491.

In evaluating these differences dealing with the implicit weighting effect, it should be noted that the average difference between the spread with quality and the spread with price was 1.4 when nonnormalized quality and price beliefs were used, and 2.1 when nonnormalized quality and objective price beliefs were used. The Pearson product-moment correlation coefficient between 1) the difference between a) the spread of nonnormalized quality beliefs measured by the difference between the brand rated lowest in quality and the brand rated highest in quality, and b) the spread of nonnormalized price beliefs measured in the same manner, and 2) tradeoff weights was .33.

If all other factors are held constant, the difference in the average correlation between measured and uniform weights was only somewhat larger when tradeoff weights were used than when direct weights were used (.0617). However, the average correlation with tradeoff weights was a good deal greater than the correlation with direct weights when nonnormalized quality beliefs and objective price beliefs were used (.1571). Since the mean difference between direct and tradeoff weights was reasonably large (.158), it appears that the average correlation with the other three versions of the model was not particularly sensitive to the differences between direct and tradeoff weights. In explaining the reasons for the lack of sensitivity, it is again necessary to mention that the vast majority of respondents stated that they were more likely of purchasing the higher quality automobiles. Although the tradeoff weight for quality was greater than the direct weight for quality with 68 percent of the respondents, the direct weight for quality was greater than or equal to .670 percent of the time that tradeoff weights were greater than direct weights. In this situation, direct weights would predict that the higher quality cars would be most likely of being purchased unless the price spread was greater than the quality spread. The price spread was greater only when nonnormalized quality beliefs and objective price beliefs were used.

Finally, since the mean correlation coefficient between

⁴Although the use of normalized beliefs might minimize the implicit weighting effect in some instances, the use of nonnormalized quality beliefs and objective price beliefs minimized it here because the vast majority of respondents seemed to attach greater weight to quality than price.

nonnormalized quality and nonnormalized price beliefs was -.8442, differences in the average correlation between various versions of the model were not suppressed by a positive correlation between beliefs.

Summary And Conclusions

The results of this study indicate that when the correlation between stated and generated attitudes is used as a criterion of validity, it is more appropriate to use measured attribute weights than uniform weights. Although this difference between measured and uniform weights might have been greater than differences in previous studies partly because of situational differences, it is doubtful that these completely explain the differences among studies. The differences also may be due to 1) a reduction in the positive correlation between beliefs 2) a reduction in the implicit weighting effect, and 3) the use of a new approach for measuring attribute importance. Although it is likely that the positive correlation and the implicit weighting effect were substantially greater in many of the previous studies, the results of this study suggest that even moderate differences among studies in terms of the correlation between beliefs and the implicit weighting effect can substantially alter the results.

Although the procedures used to reduce the positive correlation between beliefs and the implicit weighting effect were successful in this study, there are instances in which these procedures should not be used, or in which they will not be successful.

Consumers may perceive a positive correlation between beliefs based upon an objective analysis. Although this author made the attempt to select brands, evaluative criteria, and a product in which this situation would not exist, there may be good reason to investigate attitude formation when the correlation exists. It is recommended that an approach such as examining the extent that an attitude model correctly predicts attitude change be used in this situation (Bettman, Capon and Lutz, 1975).

With respect to an implicit weighting effect, the use of various measures of beliefs to alter the relative spread of beliefs across attributes might not be successful. One approach used was equalizing the spread across attributes. Yet, as mentioned previously this approach might have suppressed the difference in the average correlation between tradeoff and direct weights. The use of nonnormalized quality beliefs and objective price beliefs reduced the implicit weighting effect only because there was a high degree of respondent homogeneity in terms of the relative importance of quality and price.

The issue of whether the difference between measured and uniform weights found in this study is atypical is worthy of discussion as it relates to the external validity of these results. The difference between tradeoff and uniform weights was .236 and .163 between direct and tradeoff weights. Since the majority of the respondents attached more importance to quality than price, these differences might have been reduced substantially if respondents actually had to spend their own money. Yet, price might have become substantially more important than quality in this situation rather than the weights approaching uniformity. In addition, it is possible that respondents may have utilized rather extreme weights to allow them to differentiate clearly among objects. More specifically, since the price differences between objects tended to be similar to the quality differences, the use of uniform weights would have resulted in all objects being virtually equivalent in terms of the probability of purchase. On the other hand, one might argue that there will tend to be a dif-

ference between measured and uniform weights when respondents are allowed to select the brands used for the evaluation after being told that they should select those brands that they actually consider in making a purchase. If respondents make a selection among brands that tend to differ on one or two attributes (e.g., evaluate only subcompact automobiles or steel belted radials), one might expect that the weight of this attribute or attributes would be a good deal higher than the weight of other attributes.

Finally, since the results of this study indicate that there is a possibility that measured weights may be quite superior to uniform weights, it is relevant to ask whether there are any disadvantages in using measured weights rather than uniform weights. Other than eliminating the need to ask a single question or a limited number of questions to ascertain measured weights, which does not seem to be a very significant disadvantage, there appear to be no disadvantages in using measured weights rather than uniform weights.

References

- Frank M. Bass and William L. Wilkie, "A Comparative Analysis of Attitudinal Prediction of Brand Preference," Journal of Marketing Research, 10, (August 1973) 262-69.
- Neil E. Beckwith and Donald R. Lehmann, "The Importance of Differential Weights in Multiple Attribute Models of Consumer Attitude," Journal of Marketing Research, 10, (May 1973) 141-45.
- James R. Bettman, Noel Capon, and Richard J. Lutz, "Multi-Attribute Measurement Models and Multi-Attribute Attitude Theory: A Test of Construct Validity," Journal of Consumer Research, 1, (March 1975) 1-15.
- Joel B. Cohen, Martin Fishbein and Olli T. Ahtola, "The Nature and Uses of Expectancy-Value Models in Consumer Attitude Research," Journal of Marketing Research, 9, (November 1972) 456-460.
- Bernard Dubois, "Ideal Point Versus Attribute Models of Brand Preference: A Comparison of Predictive Validity," Advances in Consumer Research, 2, 1974, 321-333.
- Michael J. Houston and Samuel M. Gillespie "Issues in the Use of Expectancy Models of the Attitude-Behavior Relationship: An Empirical Examination," Proceedings. Southern Marketing Association, 1974, 177-79.
- Morris B. Holbrook and James M. Hulbert, "Multi-Attribute Models: A Comparative Analysis," Advances in Consumer Research, 2, 1974, 375-388.
- David G. Hughes and Jose L. Guerrero, "Testing Cognitive Models Through Computer-Controlled Experiments," Journal of Marketing Research, 8, (August 1971) 291-97.
- Michael B. Mazis, Olli T. Ahtola, and R. Eugene Klippel, "A Comparison of Four Multi-Attribute Models in the Prediction of Consumer Attitudes," Journal of Consumer Research, 2, (June 1975) 738-52.
- Dov Pekelman and Subrata Sen, "Mathematical Programming Models for the Determination of Attribute Weights," Management Science, 20, (April 1974) 1217-1229.
- Jerome E. Scott and Peter Wright, "Modeling an Organizational Buyers Product Evaluation Strategy: Validity and Procedural Considerations," Journal of Marketing Research, 13, (August 1976) 211-14.
- Jagdish N. Sheth and W. Wayne Talarzyk, "Perceived Instrumentality and Value Importance as Determinants of Attitudes," Journal of Marketing Research, 9, (February 1972) 6-9.
- Albert R. Wildt and Albert V. Bruno, "The Preference for Capital Equipment using Linear Attitude Models," Journal of Marketing Research, 11, (May 1974) 203-205.
- William Wilkie and Rolf P. Weinrich, "Effects of the Number and Type of Attributes Included on an Attitude Model: More is Not Better," Proceedings. Third Annual Conference, Association for Consumer Research, 1972, 325-340.

ASSESSING THE IMPACT OF INCREASED
PRODUCT SAFETY ON CONSUMER UTILITY

Stephen C. Burnett (student), Indiana University

Abstract

One of the costs of product safety regulation may be its effect on consumer utility. The paper presents a pilot application of conjoint analysis as a methodology for measuring the consumer utility impact of proposed safety regulation. Pending government decisions on minimum wet braking standards for bicycles and flame-resistant standards for clothing were selected for the analysis.

Introduction

Since the creation of the U.S. Consumer Product Safety Commission (CPSC) in 1972, there has been an increasing awareness among marketers that growing public and private concern over product safety will have significant implications for marketing programs. This raises questions about the probable effects of product safety regulation and how marketing should respond. Some of the suggestions offered include incorporating safety testing in new product planning (Trombetta and Wilson, 1975), product safety as a competitive strategy (Edwards, 1975; Loudenback and Goebel, 1974), and more stringent labeling and record-keeping procedures (Jensen, Mazze, and Stern, 1973). In addition to assessing the marketing implications of product safety, there has also been an increasing interest in how consumer research can assist in product safety policy decisions (Staelin and Pittle, 1977; Miller and Parasuraman, 1974) and public policy decision-making in general (Day, 1976; Wilkie and Gardner, 1974).

While consumer research in the product safety area has not been overly abundant and its focus has been somewhat fragmented, as would be expected at this early stage, the research issues emanating from the regulatory role of the CPSC are becoming more apparent. For products within its authority, the Commission has several policy alternatives including: (1) mandatory safety regulation, (2) voluntary industry standards, and (3) consumer information/education campaigns (CPSC, 1976a). Although all of these alternatives require some knowledge of consumers' behavior in order to be evaluated for effectiveness, the specific research issues raised by each are quite different.

It has been reported that only 20% of the total incidence of consumer product-related injury and death is addressable by improved product design (Staelin and Pittle, 1977). The remainder is a consequence of consumers carelessly using relatively safe products. Because of this fact, consumer information and education has become a highly attractive policy alternative and one that has begun to benefit from ongoing consumer research. A continuing project reported by Staelin (1977), Staelin and Weinstein (1974), and Thompson (1974), has concentrated primarily on whether information and education can enhance safer consumer behavior in using products. Although the researchers have found that knowledge of safety principles correlates positively with safer product use behavior, the direction of causality between knowledge and behavior remains unresolved.

A second policy alternative available to the Commission is to encourage and assist industry trade associations in the development of voluntary minimum safety standards. Despite their regulatory cost appeal, benefits of voluntary standards are obviously dependent upon the degree

to which they are implemented. Whether improved safety standards will be incorporated in products without government mandate raises the question of safety marketability. The previously cited study by Staelin and Weinstein (1974) attempted to gauge the importance of safety in purchase decisions. Their findings suggest that either because of the demand characteristics or the multidimensional nature of safety, its saliency as a product attribute varies with how its importance is measured. Using open-ended questioning, safety was found not to be an influential decision variable. However, when subjects were asked to rate the importance of safety as one of eight predetermined purchase variables, it took on a more central role in the respondents' purchase decisions. A similar result has been reported by Kuehl and Simon (1973).

Objective

The final CPSC policy alternative and the focus of the present study is to set minimum safety standards and require compliance under the threat of fines and even criminal prosecution. While this action eliminates much of the uncertainty in the actual delivery of safer products to consumers, its evaluation requires an estimation of the costs and benefits expected to accrue to consumers from a Mandatory Safety Regulation (MSR).

Consumer benefits from an MSR may take several forms, but primarily these are reduced probabilities of injury and death. There are numerous components of MSR costs. From a macro-economic perspective, some of these costs are fairly tangible such as additional resources to produce the product and expenditures for monitoring and enforcing the standard. Other cost components such as industry disruptions and increased industry concentration are more difficult to quantify.

Another cost of an MSR may be its net effect on consumer utility. As used here, utility refers simply to the total value a consumer derives from the consumption of a good or service. The concept of utility has its origins in economics, but with the growing popularity of multidimensional scaling techniques, it has more recently been applied to marketing decisions (Green and Wind, 1975; Green and Wind, 1973; Green and Rao, 1971). For a particular MSR, there may be several potential sources of utility impact. One of these is generally a price increase although an increase in price may be only one of many utility tradeoffs a consumer must make in order to obtain the utility of enhanced safety. The proposed MSR for flame-resistant treatment of clothing provides a good illustration.

The CPSC estimates that an MSR requiring all consumer apparel to meet proposed minimum flame-resistant standards would require an average price increase of 25%. In addition, the following garment performance characteristics would be affected: (1) a reduction in the wash and wear properties of some garments, (2) a reduction in the variety of fabrics available and wearer comfort, (3) a decrease in the average life of treated fabrics, and (4) some limitations on styling variations (CPSC, 1976c). Assuming that consumers derive utility from any of these attributes, then this particular MSR may have a utility cost if the increase in utility from having flame-

resistant clothing fails to offset the disutility from decreases in the performance of other product characteristics.

It is useful to note that the question of trading off utility as posed here is not concerned with whether a safety feature or performance standard is an important purchase decision variable in an absolute sense. For example, the degree to which a bicycle has the capacity to stop under wet conditions may have only minimal influence on most bike purchase decisions. Nevertheless, if utility derived from improved braking at least offsets the disutility of the price increment necessary to provide better braking, then there is no net utility cost to the consumer even though color, style, and retail outlet are the major purchase determinants. The objective of the present study was to explore the problems and potential of conjoint analysis as a technique for assessing such net utility costs. Pending CPSC decisions on minimum wet braking standards for bicycles and flame-resistant standards for clothing were selected for the application.

The idea of measuring the utility cost of government regulation is not new. Walker, Sauter, and Ford (1974) provided a conceptual framework, and Walker and Sauter (1974) used Thurstone's Law of Comparative Judgement to interally scale consumer preference for alternative retail credit contracts. The researchers' hypothesis was that regulation reducing consumer credit interest rates may cause retailers to compensate for lower interest charges by requiring higher down payments and monthly installments plus increasing the price of the merchandise. When these strategic "secondary effects" are taken into account, total consumer utility may be decreased by the credit regulation.

In terms of the type of public policy evaluation issue addressed, the present study parallels the works cited above. There are, however, two distinctions: one with respect to the requirements of the policy area, the other with respect to methodology. For consumer credit regulation, utility tradeoffs result from strategic responses of retailers. Because of the uncertainty concerning what these changes would be, Walker and Sauter (1974) had to rely on rough estimates. Given this constraint, their conclusions hold only to the extent that these approximations reflect how retailers would actually respond. For product safety decisions, there fortunately is somewhat less imprecision in determining both the product characteristics that would be affected and by what amounts. Product modifications are conducive to the use of engineering tests and prototypes which provide relatively accurate estimates of performance tradeoffs.

The second difference between the two studies concerns the flexibility of Thurstone scaling versus conjoint measurement. Both scaling models may require identical subject tasks when scaling the average utility for a group of consumers. In fact, Curry and Rogers (1977) recommend using the two methods in tandem when aggregating responses across subjects. If there is an interest in utility at the individual subject level, however, the Thurstone model requires each stimulus or product concept to be presented to a subject a large number of times (Torgerson, 1958). With conjoint measurement the scaling of individual respondents is much easier since the stimuli must be judged only once. Because individual subject comparisons were important to the present study, conjoint measurement was the preferred technique.

Study Design

Since the purpose of the study was exploratory, a convenience sample of 40 undergraduate students was used. This made it necessary to select products with which students are familiar. Ten-speed bicycles and shirts or

blouses were thought to meet this criterion fairly well. Also, a preliminary questionnaire indicated that two-thirds of the subjects owned 10-speed bicycles. As previously noted, the flame-resistant standard now pending before the CPSC would affect a garment's price, durability, and easy-care characteristics (CPSC, 1976c). The only impact of the proposed wet braking standard for bicycles is an estimated price increase of 10% (CPSC, 1976b).

Factor Descriptions and Level Specification

The general procedure in conjoint analysis is to have subjects judge in terms of their preference or some other criterion a series of product descriptions or scenarios. The products systematically differ on several attributes (factors) by specific amounts (levels). With a subject's preference ranks as input, the conjoint measurement algorithm derives interally scaled utility weights (part-worths) for each factor level. With an additive scaling model, such as Kruskal's MONANOVA used in this study, the sum of the part-worths must preserve the original rank orders (Green and Wind, 1973).¹ To determine the degree to which a subject's utility is altered by paying \$150 rather than \$165 for a particular product, one only needs to examine the absolute change in part-worths for the two prices relative to the changes in part-worths of the other factor levels. The greater the relative utility change, the greater the utility attached to the product attribute.

A critical design aspect of conjoint analysis is the specification of the factor levels. The relative utility of a factor is a direct function of its levels. For example, the part-worth of a lower price could be increased dramatically by selecting polar price levels such as \$100 and \$300 as opposed to \$100 and \$120. Thus, when scaling the utility of a proposed safety feature versus the utility of tradeoff attributes, it is crucial that objective criteria exist for selecting the factors and their levels.

For the present study, limited CPSC test reports were available to guide factor level specification. The only quantitative estimates published by the Commission were anticipated price increases for the two safety feature proposals. For this reason, factor levels were set based on data from recent Consumer Reports tests of 10-speed bicycles and men's shirts. Because of this limitation, the relative utilities of the factors reported later in the paper should be interpreted solely as an illustration of the informational value of conjoint analysis. Different factor levels or different ways to operationalize the factors could produce different findings from those obtained here.

Exhibit 1 lists brief descriptions of the bicycle factors and their levels. For both shirts and bicycles, four factors at two levels were used for a 2⁴ full factorial design. Since the wet braking standard causes only a price tradeoff, in theory braking and price factors would have sufficed. However, because prior studies (Kuehl and Simon, 1973; Staelin and Weinstein, 1974) suggest that the topic of safety may be subject to demand characteristics, two additional moot factors were used in conjunction with price and braking. This was done to insure that respondents would be unaware of the study's purpose and to make the task more realistic.

¹MONANOVA is a popular nonmetric model and has been found to be appropriate in a variety of studies. It should be noted that other methods are also suitable for analyzing conjoint data (e.g., see Cattin and Wittink, 1977).

EXHIBIT 1

Bicycle Factor Descriptions and Levels

<u>Form A Factors</u>	<u>Levels</u>
(1) <u>Price</u> --refers to the retail price of the bike rounded to the nearest dollar	\$150 and \$165
(2) <u>Overall Pedaling Ease</u> --refers to the overall reward the rider receives from his or her pedaling efforts	rated Fair and Very Good
(3) <u>Wet Braking Ability</u> --refers to the maximum number of feet required to stop from 15 mph with wet wheels	20 and 50 feet
(4) <u>Frame Quality</u> --refers to whether major frame joints were reinforced and how well fork ends were attached to the frame	rated Fair and Very Good

<u>Form B Factors</u>	<u>Levels</u>
(1) <u>Price</u> --refers to the retail price of the bike rounded to the nearest dollar	\$150 and \$165
(2) <u>Handling Precision</u> --refers to how obedient the bike was and how much rider attention it demanded	rated Fair and Very Good
(3) <u>Wet Braking Ability</u> --refers to the maximum number of feet required to stop from 15 mph with wet wheels	20 and 50 feet
(4) <u>Shifting Ease and Precision</u> --refers to how certain the shifting mechanism was at arriving at the right gear and how much rider attention and effort was required to find the desired gear	rated Fair and Very Good

The hypothesis to be tested was that Form B subjects would attach higher utility to a flame-resistant feature than Form A subjects because of greater knowledge of clothing fire hazards.

This hypothesis and its conceptual underpinnings are far too simplistic to yield substantive conclusions. Knowledge of a product's hazard potential could be manipulated on several dimensions such as the probability of injury to a specific subject or to people in general. Moreover, both the probability of an injury and its severity could influence safety feature preference to different degrees. The strength of the statement could also have implications for utility. The study's mild manipulation of knowledge was included in the design primarily to illustrate the technique's potential for experimentation with safety knowledge. Complete descriptions of the shirt/blouse factors and levels are shown in Exhibit 2.

EXHIBIT 2

Shirt/Blouse Factor Descriptions and Levels

(1) <u>Price</u> --refers to the average retail price of the garment rounded to the nearest dollar	\$12.00 and \$15.00
(2) <u>Durability</u> --refers to the average number of years the garment should last under normal use and suggested laundry procedures	1½ and 2 years
(3) <u>Flame-Resistant Characteristics</u> --refers to whether or not the garment would actually ignite or catch fire when exposed to an open flame	yes and no
(4) <u>Wash and Wear Properties</u> --refers to whether or not the shirt or blouse retained its original shape and did not require ironing after proper washing and tumble-drying	yes and no

The inclusion of moot factors raises the important methodological question of whether different sets of moot factors will alter the relative utilities of the price and braking factors. To address this question, subjects were randomly assigned to one of two groups (labeled Forms A and B). Form A subjects received overall pedaling ease and frame quality as moot factors while Form B subjects were given handling precision and shifting ease. Although any bicycle attributes would have served equally well, all four of the moot factors were characteristics of bicycles tested by Consumer Reports.

Before discussing the subject's task, one other aspect of the factor descriptions should be mentioned. Past research on safety knowledge and safer product use behavior has not explicitly dealt with the effect of safety knowledge on product choice behavior. To explore how conjoint analysis might be helpful in examining a link between a consumer's knowledge of a product's hazards and preference for a safety feature, subjects were also randomly divided into two groups for the shirt/blouse task (again labeled Forms A and B). Both groups were told in the instructions that a garment was rated as flame-resistant if it would not ignite when exposed to an open flame. The Form B instructions, however, contained the following introductory sentence to the flame-resistant definition:

"Because some 28,000 people are burned annually from clothing fires, Consumer Reports tested each garment for flame-resistant characteristics."

Subjects' Task

The task required of the subjects was divided into two sections. They were first asked to rank in terms of their preference all 16 product combinations. To simplify the ranking task, subjects were instructed to sort the products into high and low preference piles, to rank order each pile, and then to finally combine both piles into one strict rank. The subjects' second task was to provide self-explicated factor importance weights. This was accomplished with a constant sum scale. They were asked to divide 100 points among the factors such that the point assignments indicated the relative degree to which each factor and its levels influenced their product preference.

The reason for obtaining self-explicated factor importance weights was to check for convergent validity. Deriving the relative importance of product attributes from rank orders of product scenarios is a distinctly different measurement method from having subjects directly report this importance. The extent to which the two measures correlate should be indicative of convergent validity. The instructions and tasks for bicycles and garments were identical.

Analysis of Results

The initial step of the data analysis was to identify subjects who either could not perform the ranking task or whose decision rules were inappropriate for MONANOVA's additive composition rule. The criteria for identifying and eliminating these subjects from additional analysis were stress values and factor directionality. Stress is the badness-of-fit measure of the scaling model to the original data. Factor directionality refers to whether a subject's utility increases or decreases in the hypothesized manner with changes in a factor's levels. For both products, an increase in price was hypothesized to lower utility while increased utility was expected for the higher levels of all other factors.

Of the total number of rank orders submitted to MONANOVA ($n = 35$ for bikes, $n = 40$ for garments), only three subjects exhibited stress values greater than the .15 cut-off point. Five subjects violated hypothesized directionality. The low percentage of subjects eliminated is evidence that the subjects could perform the task as requested and that the additive composition rule of MONANOVA was adequate. In fact, 82% of the total rank orders had stress values of less than .01.

The means and standard deviations of the MONANOVA and self-explicated factor importance weights are shown in Table 1. The method for obtaining self-explicated weights has been discussed, but the MONANOVA factor importance weights require explanation. Based on a subject's 16 rank orders, the output of the model is eight intervally scaled utility values--one for each factor level. Since these part-worths comprise an interval scale with unique origin and unit for each subject, it is not permissible to compare part-worths across subjects. It is necessary therefore, to convert the part-worths for each subject into four ratio scaled factor importance weights (FIW). These are computed by simply dividing the absolute value of the change in the two part-worths of each factor by the total absolute value of changes in part-worths for all four factors. Factor importance weights may be interpreted as the percentage contribution of a particular factor to a subject's total utility for all factors.² As an illustration, if subject X's FIW for price is 25 compared to 50 for subject Y, then Y may be said to attach twice as much utility to the lower price than X. Also, Y's utility for the lower price is equal to his combined utility for all other factors.

From Table 1, it is apparent that the standard deviations of the MONANOVA factor weights were high relative to the mean values. This is a strong indication that it may be inappropriate to compute an average group scale since subjects tended to have quite different utility functions. Because the sample was demographically homogeneous, the high variance may suggest the existence of utility profile segments resulting perhaps from non-demographic variables. An extension of the study, which would help better understand safety feature utility, would be to form clusters of subjects with similar factor importance weights. These clusters could then be examined for systematic differences in such variables as attitudes toward safety, product experience, and safety knowledge and behavior.

Another interesting finding is evident when the means of the self-explicated weights are contrasted with the MONANOVA weight means. For bicycle Forms A and B separately

and the combined shirts/blouses forms,³ the self-explicated price factor weights were greater than the MONANOVA price weights. Similarly, subjects tended to overestimate the importance of the durability factor for garments compared with the importance of durability as reflected by their preference ranking behavior. It is not completely apparent why price and durability were overestimated. One could perhaps speculate that price and durability are common and rational product attributes which may not influence product preference to the extent that self-report measures would imply. Despite these differences, the product moment correlations for the MONANOVA and self-explicated weights shown in Table 2 were statistically significant ($p < .05$) for all factors. This agreement is evidence of the convergent validity of the two measures.

One last aspect of the individual MONANOVA factor importance weights should be noted. For bicycle Form A and B groups, the utility of improved wet braking more than compensated for the disutility of the 10% price increase when averaged over subjects. Conversely, the utility of the flame-resistant feature was far outweighed by the utility of price, durability, and wash and wear properties. This finding is consistent with the CPSC's opinion that consumers probably would not be willing to forego the benefits of these attributes for a flame-resistant feature (CPSC, 1976c). Again, these findings should be regarded solely as illustrations of the methodology because of the high degree of subjectivity in the factor level specification.

The question of whether different sets of moot factors would affect the relative factor importance weights of price and wet braking was examined by using analysis of variance. The criterion variable was the ratio of the price and wet braking factor weights. Form A or B, bike ownership, and sex were specified as the independent variables. As shown in Table 3, neither main- nor interaction-effects were significant at any commonly used significance level, suggesting that moot factors can be added to make the task more realistic without affecting the stability of the relationship between the factors of interest.

Analysis of variance was also used to test the hypothesis that garment Form B subjects would have greater flame-resistant factor weights from Form A subjects because of more knowledge of clothing fire hazards (Table 4). Sex was included as an independent variable because it was thought the sex of the respondent could interact with the manipulated knowledge. The results show, however, that while the main-effects of knowledge and the interaction-effects of sex and knowledge were not significant, sex as a main-effect was highly significant ($p = .001$). The average flame-resistant factor weight for females was 23 compared with an average weight of 9 for males. An ANOVA run for the wet braking factor weights revealed that sex had no significant effect for that safety feature. These findings seem to suggest that safety utility should be treated on a product and consumer specific basis.

Conclusions

Limitations

For both public policy and marketing strategy decision-making, conjoint analysis has several limitations. Foremost among these is the artificiality of the subjects'

²Another measure of a factor's importance is the percentage of total variance it explains.

³As discussed later, garment Form A and Form B groups did not significantly differ on any of the factor importance weights and were aggregated for the bulk of the analysis.

task. Respondents are not making market place decisions but rather indicating preferences for abstract descriptions of products which differ on a small number of attributes. While this procedure may be preferable to self-report measures of attribute saliency, the question remains whether conjoint analysis conclusions will be manifested in purchase behavior. Such a limitation is not overly troublesome from the standpoint of assessing the utility impact of a proposed safety regulation. It is, however, quite pertinent for the bicycle manufacturer who must estimate the sales response to the 10% price increase which will accompany improved braking performance.

Another limitation of the technique is the possibility that a product attribute may be imputed to possess utility solely as a consequence of its inclusion as a factor. A consumer may have never considered a garment's flame-resistant properties, yet may feel it must somehow be important since the researchers selected it for the task.

The potential for distorted conclusions from inaccurate factor level specification has been repeatedly emphasized. A second source of error which could cause inappropriate conclusions is the reliability of the scaling technique. The time frame of the study prohibited a check of test-retest reliability. However, the finding that garment Form A and B groups displayed no significant differences on any of the factor importance weights is weak evidence of reliability. Before conjoint analysis is used for conclusive research, reliability should receive additional attention.

Extensions

Throughout the report several extensions of this study's very basic application of conjoint measurement have been recommended. In addition, the technique has the potential for assisting policy makers in identifying the utility costs for specific products within a product category. The finding that a student's utility for a shirt or blouse would be substantially reduced by a flame-resistant feature is hardly surprising. But would the same conclusion be reached for pajamas and robes? Or would the utility impact be the same for elderly consumers where the probability of clothing fires are higher? With larger, more heterogeneous samples and multiple products within product categories, conjoint analysis has the capability for answering these questions. Depending on the utility costs, it could be that a clothing flammability standard should not apply to all types of apparel, but instead only to those where consumers are more willing to make the necessary utility tradeoffs.

Product safety regulation has been justified on the basis that consumers are unable to make tradeoffs between the benefits derived from using a product and the hazards this use may entail (Jones, 1973). By combining existing CPSC information with findings from conjoint analysis applications, this question can be empirically tested. From its national information system, the CPSC has product specific estimates of injury probabilities for various socio-economic classes of consumers. The degree to which consumers with different injury probabilities for a given product have concomitant factor importance weights for a safety feature designed to reduce these hazards may be an indication of whether consumers are capable of incorporating safety in their purchase decisions.

REFERENCES

Philippe Cattin and Dick R. Wittink, "Further Beyond Conjoint Measurement: Toward A Comparison of Methods," in William D. Perreault, Jr., (ed.), Advances in Consumer Research, Vol. IV, 1977, 41-45.

David Curry and William Rodgers, "Aggregating Responses in Additive Conjoint Measurement," in William D. Perreault, Jr., (ed.), Advances in Consumer Research, Vol. IV, 1977, 35-40.

Ralph L. Day, "Prescription for the Marketplace--Everyone Listen Better," Business Horizons, 19 (December 1976), 57-64.

Alfred L. Edwards, "Consumer Product Safety: Challenge for Business," University of Michigan Business Review, 27 (March 1975), 18-22.

Paul E. Green and Vithala R. Rao, "Conjoint Measurement for Quantifying Judgmental Data," Journal of Marketing Research, 8 (August 1971), 355-63.

Paul E. Green and Yoram Wind, Multiattribute Decisions in Marketing: A Measurement Approach, (Hinsdale: The Dryden Press, 1973).

Paul E. Green and Yoram Wind, "New Way to Measure Consumers' Judgements," Harvard Business Review, 53 (July-August 1975), 107-117.

Walter Jensen, Jr., Edward M. Mazze, and Duke N. Stern, "The Consumer Product Safety Act: A Special Case in Consumerism," Journal of Marketing, 37 (October 1973), 68-71.

Mary Gardiner Jones, "Social Responsibility: The Regulator's View," California Management Review, 15 (Summer 1973), 78-84.

P. G. Kuehl and M. E. Simon, "The FDA Listens: A Survey of Consumer Opinions," FDA Consumer, (March 1973), 8-13.

Lynn J. Loudonback and John W. Goebel, "Marketing in the Age of Strict Liability," Journal of Marketing, 38 (January 1974), 62-66.

Joseph C. Miller and A. Parasuraman, "Advising Consumers on Safer Product Use: The Information Role of the New Consumer Product Safety Commission," in Ronald C. Curhan, (ed.), American Marketing Association 1974 Combined Proceedings, 1975, 372-376.

Richard Staelin, "The Effects of Consumer Education on Consumer Product Safety Behavior," in William D. Perreault, Jr., (ed.), Advances in Consumer Research, Vol. IV, 1977, 380-387.

Richard Staelin and David Pittle, "Consumer Product Safety: A Discussion Paper," American Marketing Association Consumerism Workshop Monograph, (Chicago: A.M.A. 1977).

Richard Staelin and Alan G. Weinstein, "Correlates of Consumer Safety Behavior," in Scott Ward and Peter Wright, (eds.), Advances in Consumer Research, Vol. I, 1974, 87-100.

Jeffery L. Thompson, "Product Safety: Suggestions for Better Use and Purchase Behavior Through Consumer Education and Information," in Scott Ward and Peter Wright, (eds.), Advances in Consumer Research, Vol. I, 1974, 101-107.

Warren S. Torgerson, Theory and Methods of Scaling, (New York: John Wiley and Sons, 1958).

William L. Trombetta and Timothy L. Wilson, "Foreseeability of Misuse and Abnormal Use of Products by the Consumer," Journal of Marketing, 39 (July 1975), 48-55.

U.S. Consumer Product Safety Commission, Annual Report, Fiscal Year 1976, (Washington: U.S. Government Printing Office, 1976)a.

U.S. Consumer Product Safety Commission, "Product Profiles: Bicycles," October 1976, b.

U.S. Consumer Product Safety Commission, "Product Profiles: Wearing Apparel," October 1976, c.

Orville C. Walker, Jr. and Richard F. Sauter, "Consumer Preferences for Alternative Retail Credit Terms: A Concept Test of the Effects of Consumer Legislation," Journal of Marketing Research, 11 (February 1974), 70-8.

Orville C. Walker, Jr., Richard F. Sauter, and Neil M. Ford, "The Potential Secondary Effects of Consumer Legislation: A Conceptual Framework," Journal of Consumer Affairs, 8 (Summer 1974), 144-155.

William L. Wilkie and David M. Gardner, "The Role of Marketing Research in Public Policy Decision Making," Journal of Marketing, 38 (January 1974), 38-47.

TABLE 1

MONANOVA and Self-Explicated Factor Importance Weights

	Individual MONANOVA Factor Importance Weights		Self-Explicated Factor Importance Weights	
	Mean	Std. Dev.	Mean	Std. Dev.
<u>Bikes--Form A Factors (n=17)</u>				
Price	17.35	15.49	27.41	18.49
Pedaling Ease	27.88	12.77	25.41	9.50
Wet Braking	24.11	22.07	21.82	14.79
Frame Quality	30.47	17.46	25.35	13.29
	100.00		100.00	
<u>Bikes--Form B Factors (n=15)</u>				
Price	8.80	10.13	16.33	12.45
Handling Precision	36.27	22.37	33.33	13.32
Wet Braking	28.93	25.69	25.66	12.37
Shifting Ease	26.00	17.41	24.67	10.77
	100.00		100.00	
<u>Shirts/Blouses-- Both Forms (n=35)</u>				
Price	17.57	18.33	27.34	9.89
Durability	16.42	14.36	24.88	8.91
Flame-Resistance	14.97	12.72	11.22	7.77
Wash and Wear	50.48	22.03	36.54	14.78
	100.00		100.00	

TABLE 2

Correlation Coefficients
MONANOVA Weights vs. Self-Explicated Weights

<u>Factor</u>	<u>Bicycles--Form A (n=17)</u>	
	<u>Coefficient</u>	<u>Significance^a</u>
Price	.78	.001
Pedaling Ease	.50	.020
Wet Braking	.55	.010
Frame Quality	.65	.002

<u>Factor</u>	<u>Bicycles--Form B (n=15)</u>	
	<u>Coefficient</u>	<u>Significance</u>
Price	.46	.040
Handling Precision	.81	.001
Wet Braking	.62	.006
Shifting Ease	.85	.001

<u>Factor</u>	<u>Shirts/Blouses--Both Forms (n=35)</u>	
	<u>Coefficient</u>	<u>Significance</u>
Price	.43	.005
Durability	.53	.001
Flame-Resistance	.59	.001
Wash and Wear	.45	.004

^aSignificance Level = one-tailed t-test.

TABLE 3

Analysis of Variance--Bikes
(dependent variable = ratio of price and braking factor importance weights)^a

<u>Sources of Variation</u>	<u>Sums of Squares</u>	<u>DF</u>	<u>Mean Square</u>	<u>F-Ratio</u>	<u>Significance</u>
Main Effects:	613	3	204	1.011	.405
Form A or B	86	1		.427	.520
Bike Ownership	231	1		1.145	.295
Sex	207	1		1.023	.322
Two-Way Interactions:	591	3	197	.975	.421
Form/Bike Ownership	1	1		.007	.933
Form/Sex	295	1		1.459	.239
Sex/Bike Ownership	255	1		1.261	.273
Three-Way Interactions:					
Form/Bike Ownership/Sex	9	1		.046	.831
Total Explained	1214	7	173	.858	.533
Residual	4853	24	202		
Total Variation	6067	31	196		

^aComputed by MONANOVA

TABLE 4

Analysis of Variance--Shirts/Blouses
 (dependent variable = flame-resistant factor importance weights)^a

<u>Sources of Variation</u>	<u>Sums of Squares</u>	<u>DF</u>	<u>Mean Square</u>	<u>F-Ratio</u>	<u>Significance</u>
Main Effects	1741	2	870	7.207	.003
Form A or B (Knowledge)	104	1		.863	.360
Sex	1741	1		14.413	.001
Two-Way Interaction:					
Form/Sex	25	1		.206	.653
Total Explained	1765	3	588	4.873	.007
Residual	3743	31	121		
Total	5508	34	162		

^aComputed by MONANOVA

TRANSPORTATION ATTITUDES OVER TIME:
A LONGITUDINAL APPROACH¹

Mark I. Alpert and Linda L. Golden
The University of Texas at Austin

Abstract

Research into consumer transportation modal choice attitudes is receiving increased emphasis. This paper compares public modal choice criteria reflected at the peak of the "energy crisis" with those obtained one year later. Despite a relaxing of gasoline lines, no significant changes in ecological variables were noted, and other travel criteria remained constant. Public transportation remains perceptually inferior to private, and resistance to several mass transit subsidy plans may be increasing.

Marketing and Consumer Psychology have recently begun to find application in attempts to model and promote mass transportation usage and public support (Blattberg and Stivers, 1970; Business Week, 1974; Hille and Von Cube, 1963; Hille and Martin, 1967; Mundy, Cravens and Woodruff, 1974; Sheth, 1975). Many of these studies have stressed the need to identify attitudes relevant to modal choice decisions and then tailor a marketing mix toward target markets of potential riders (and, occasionally, toward non-riding supporters of the mass transit system). However, little has been done to test the reliability of the scales that have been used to measure these attitudes toward transportation and its features. Techniques that have been developed elsewhere have been employed in this problem area, thinking that validity and reliability can be generalized across products. This may be true, but it would be useful to know if the attitudes that have been measured are relatively stable over time. Modifying an existing transportation system to build in benefits revealed as important deficiencies at one point in time might take so long that by the time the political and system problems have been negotiated (at great expense—e.g., BART), public attitudes may have changed and anticipated support may not develop. Therefore, we need to know whether the attitudes revealed in one study are still salient when measured again, say, a year later. If so, planning may be more secure. If not, one might extrapolate from longitudinal changes what will be the trends in attitudes and relative importance of modal choice criteria, so that these changes in the "ideal system" may be properly anticipated and programmed.

In the spring of 1974, when most of this study's first year's data were collected, the "energy crisis" was first being widely perceived by the public, with lengthened gasoline station lines, rapidly rising prices for gasoline, and increased rhetoric about the Arab oil embargo, self-sufficiency, and related issues. By the time Spring of 1975 had come, gas lines were a (temporary?) thing of the past, and talk of gasoline rationing was not heard amongst the general public. Accordingly, it seemed appropriate to explore whether the general public in the survey area might still seek the same configuration of transportation features, including the relatively high determinance of energy savings and low pollution per passenger. Comparisons could also be made between the relative desirability of various funding proposals for public transportation, given changes in economic and political circumstances of these two years. Changes in perceived images of private autos, buses, and the differences between the two, could also be monitored to see whether public transit was perceived as "gaining" significantly, due to changes in public attitudes transit improvements, or some combination.

Clearly, studying trends in attitudes and transit priorities over a two year period in one study area does not constitute a sufficient data base for generalizing about long-term trends. However, it was felt that some insight into sensitivities to "the energy crisis," and possible changes in a variety of criteria for modal choice might be gained through examining a community's responses to the same questions one year "after the crisis". To the extent that results can be generalized to the broader United States culture, trends (or stability) in this area might be indicative of broader, transportation-relevant societal trends as well.

Method

Sample and Data Collection

The findings reported in this paper are drawn from questionnaires administered to two separate samples of adults in Austin, Texas, a medium-sized (population about 300,000) southwestern city.

Like most cities built after the automobile, there is a relatively small central city surrounded by generally low density housing and decentralized centers for shopping and major industry. The city has grown very rapidly in the last ten years, and most of this resembles the kind of urban sprawl perhaps exemplified by Los Angeles. However, there remains a central employment-shopping attraction because of the centrally located University of Texas campus (42,000 students), and state capital complex. In addition, there has been a recent surge of downtown renewal and an attempt to establish evening theater and other entertainment.

Similar to other decentralized cities, travel patterns reflect the dominance of the car. According to the city's own surveys over 97% of trips are taken in private automobile. The bus service is marginal in areas of higher usage—lower income to downtown, but service is generally sparse throughout the rest of the city, with routes along major arteries, but relatively infrequent headways especially in off peak times.

During April through June, 1974, 252 usable responses were obtained from adults who were contacted in a stratified by census-tract (quotas proportional to population) area random sample of city households. Interviewers enumerated households within each census tract, with starting points determined by the researchers selecting random blocks within tracts and random corners and walking directions within blocks. Every third household was approached, with provisions for call-backs, staggered interviewing hours, and alternately selected male and female respondents (18 years and over). Due to the length and complexity of the survey instrument, interviewers were essential in insuring cooperation and providing clarifications to questions. To increase the

¹We gratefully acknowledge the support of the Department of Transportation program of University Research, Contract DOT-OS-30093.

speed and candor of responses, respondents filled out their own questionnaires, except in those households where translation to Spanish necessitated a more active role by the bilingual interviewers.

Data were collected in the same manner one year later, except that the sample size was trimmed by about one-third (159 usable responses), and certain sections of the questionnaire (not relevant to this analysis) were modified. While a longitudinal study ideally involves obtaining data from the same persons during two or more different time periods, this approach was varied for this study. We decided to apply similar criteria for sample selection in both years, but not to attempt to interview the same households.² It was felt that persons who were willing to respond a second time to our questionnaire, after having spent 45 minutes doing so a year earlier, were likely to be more positively biased toward public transportation than were those who merely completed the process one time. A lower response rate was also likely, and this would increase survey costs while lowering reliability. Inability to contact people who had moved since the first interview would also bias results, since original respondents were contacted door-to-door, and not from a year-old list. The study, thus, is longitudinal in the sense that similarly selected samples are used to generalize parameters for the survey area in 1974 versus 1975. Given the care that was taken to choose samples representing the community adult population, it is argued that differences between the mean responses for the two survey years would represent community attitude changes over time, provided these differences were significantly greater than those that could be allowed due to random sampling fluctuations. (Some demographic differences between the two samples were observed, but as will be discussed, these had minimal impact on the key comparisons between respondents for the two survey years.)

Overview of Questionnaire and Data Analysis

Part One of the questionnaire first obtained information concerning respondents' traveling frequency for trips to work (or school, if students) and mode usually selected. Next, subjects were asked to assume they were choosing a transportation mode for trips to work or school, and to evaluate 27 modal attributes (e.g., economy, convenience, energy use per passenger...) in two distinct ways. Initially, they were to indicate the relative importance of each attribute on a five-point scale ranging from "no importance" (scored as 1) to "extremely important" (scored as 5). Then, they indicated how much difference they perceived among various transportation modes in terms of each attribute listed. Five-point rating scales were again used, ranging from "no differences" (1) to "extreme differences" (5). Scale positions had been previously found to be approximately equal intervals (Myers and Warner, 1968).

Scales for these importance and difference perceptions were multiplied together for each respective attribute and respondent, to obtain a measure of the "determinance" (Alpert, 1971; Myers and Alpert, 1968) of that attribute. Potential determinance scores for each subject and attribute thus ranged from 1 (no importance, no differences) to 25 (extremely important, extreme differences).

This combining procedure is based on the assumption that the relative weight of a specific attribute in determining whether or not a particular transportation mode is

²Walking directions from starting points were arbitrarily shifted ninety degrees to minimize the chance that the same persons might be contacted two years in a row (none were).

selected is a function of the combined effect of the importance of the attribute to travelers and the amount of perceived variation among alternative modes, in terms of that attribute. For example, avoiding traffic congestion was perceived as important, but it probably lacks determinance because many modes (in the survey area) are perceived as equally free from (or subject to) traffic congestion. Accordingly, perceptions of this attribute of local transportation modes probably do not influence modal choices nearly as much as, say, dependability, which has both importance and perceived variation among modes.

After providing these importance and difference perceptions, respondents next rated a personal car along five-point semantic differentials (e.g., Economical: _____: _____: _____: Expensive) to indicate its suitability for these commuter trips in terms of each of the above 27 attributes. This format was also used to secure ratings of a bus's attributes for the same trip purpose.

Respondents indicated their attitudes toward financing public transportation, as well as transportation's role in city planning, pollution, and so forth. This section also contained a "would you use city mass transit if improved..." question which was used to identify non-users of public transportation who would be likely switchers to an improved system. Comparisons between this segment and the rest of the sample(s) have been reported elsewhere, and provide insight into segmentation strategies for public transportation marketing. Two additional questionnaire sections obtained media and demographic data to assist in this segmentation. In this paper the demographic data are used to compare the two years' samples, to see if demographic differences across samples might influence on the divergence (or lack thereof) of findings for the different survey period.

Results

Table 1 presents a descending ranking of the determinance scores of the 27 characteristics of transportation modes for work/school trips, as rated by the 1974 sample. The "z-values" represent the comparison of the mean determinance rating for each attribute with the mean for all attributes, adjusting for the standard deviation of these ratings, and the number of persons rating each attribute. This is not a strict statistical test, since the true universe mean and sigma are unknown, but it provides a reasonable cut-off for "how high is high." The right-hand column summarizes the results of comparing the perceived images of cars vs. buses for commuter trips, in terms of attributes such as economy, dependability, and the like. (Table 4, below, gives mean scores for each).

One can note that for the ten attributes which were significantly high in determining modal choices, buses were viewed as superior in six and cars in only four. Table 4 provides statistical details and mean image profiles for these comparisons, which were analyzed using analysis-of-variance, with repeated measures (bus vs. car) for each dependent variable (Veldman, 1967). While a more precise quantification of the utility model underlying modal choices is being estimated by currently ongoing conjoint measurement research by the authors, Tables 1 and 3 may be interpreted to show that cars had sufficiently large perceived superiorities along highly determinant attributes (such as convenience and dependability) that more than offset the perceived superiorities (typically smaller in magnitude) of buses in features less determinant of modal choice. Further, since this set of features was drawn from traits generally considered to be positive benefits in transportation modes, one might argue that virtually all of

TABLE I
Determinance Scores and Modal Comparisons
for 1974 Adult Sample
(Work/School)

TABLE 2
Discrimination Between Determinance Scores
1974 versus 1975

Rank	Attribute	Z Value	Car or Bus Superior	1974	1975	F-ratio
1	Convenience	8.85 ²	Car ²	13.4	13.9	.40
2	Dependability	8.06 ²	Car ²	15.5	15.0	.41
3	Economy	5.14 ²	Bus ²	12.7	12.2	.41
4	Freedom from repairs	4.53 ²	Bus ²	7.4	7.8	.30
5	Low energy use per passenger	4.33 ²	Bus ²	10.6	10.3	.22
6	Brief travel time	4.05 ²	Car ²	6.8	6.3	.74
7	Low pollution per passenger	3.91 ²	Bus ²	10.4	10.4	.00
8	No parking problems	3.88 ²	Bus ²	4.6	4.9	.32
9	Flexibility	3.20 ²	Car ²	12.6	13.6	1.22
10	Freedom from accidents	1.68 ¹	Bus ¹	13.5	12.6	1.28
11	Ease of travel with packages	1.51	Car ²	10.6	11.0	.25
12	Uncrowded	.41	Car ²	11.4	11.6	.08
13	Freedom from weather	.39	Car ²	6.8	7.3	.54
14	Avoid traffic congestion	-.03	n.s.d.	13.0	13.7	.54
15	Safe from dangerous people	-.17	Car ²	10.2	10.9	.57
16	Relaxing	-.61	n.s.d.	12.6	12.6	.00
17	Privacy	-1.09	Car ²	10.0	9.7	.11
18	Ease of travel with children	-1.15	Car ²	11.2	10.9	.14
19	Pleasant riding surroundings	-2.98	Car ²	7.9	7.1	1.22
20	Quiet ride	-3.63	Car ²	7.0	6.2	1.03
21	Ability to look at scenery	-4.13	Bus ²	12.7	12.7	.01
22	Smooth ride	-5.06	Car ²	7.0	6.4	.63
23	Can listen to radio or tape	-5.55	Car ²	14.8	14.1	.53
24	Ability to read	-5.70	Bus ²	8.4	8.5	.00
25	Opportunity to socialize	-6.10	Bus ²	9.5	9.2	.19
26	Fun to drive	-6.16	Car ²	9.5	9.0	.28
27	Socially accepted transportation	-9.92	Car ²	8.1	8.8	.96

¹p < .05 (two-tailed test)
²p < .01 (two-tailed test)

them are desired, albeit only ten, were seen as exceptionally desirable. Across the set of 27 features, cars were seen as superior in 16, buses in nine, and two features were non-differentiating.

While specific policy recommendations cannot be made without directly analyzing the determinant attributes for potential switchers to public transit (rather than the general public), this longitudinal study sought to compare general community attitudes and criteria for modal choice. Overall changes would be important indicators of general community trends, independent of their importance to various sub-segments of transportation interest. Table 2 shows a remarkable degree of similarity between the profile of determinance scores derived during the two years. Observation of the means for each attribute for both years (averaging the product of importance x perceived differences for each attribute, within each sample), shows almost identical statistics for both years. Attempting to discriminate Year One versus Year Two respondents on the basis of these 27 variables would be futile, since the Wilks Lambda statistic evaluated by the linear discriminant analysis model has an estimated 97 percent probability of being due to chance or sampling fluctuations. In other words, one could not assert that the general profile of criteria for modal choice changed from 1974 to 1975 without taking a 97 percent chance of being incorrect. Furthermore, not one of the attributes was rated as significantly more or less determinant in 1975 than in 1974, even though at the .05 level of significance one would expect between one or two such fluctuations due to chance. Of specific interest is the fact, shown in Tables 2 and 4, that energy usage and pollution remain important criteria (and perceived advantages of public transportation), one year after the temporary peak in the "energy crisis." Freedom

Wilks Lambda = .940
Probability = .97

$\bar{x}_1 = 10.30$ $\bar{x}_2 = 10.25$
Possible range 1-25

from repairs and parking problems may be gaining, but not significantly so, and these kinds of variations have to be considered due to sample fluctuations. Should any trends develop over a longer time span, changes in determinance of various features may prove relatively favorable or unfavorable to public transportation. At this point, the relative modal choice criteria in this community seem stable and retain the mix of attributes in which public transportation was initially seen as superior in some traits and inferior in others. Next, let us examine whether changes in the relative ability of these modes to provide these features changed during this one year period.

Tables 3 to 6 provide considerable detail regarding relative images of cars versus buses during both years, as well as changes in car image and changes in bus image over time. Examining these data, one would have to conclude much the same things as were said above about criteria for modal choice. Not only were the determinance scores stable, but people's perceptions of the relative ability of buses versus cars in supplying these attributes were essentially stable during this time period. The 1974 mean profiles of car versus bus shown in Table 3 indicates patterns of relative superiority for cars in convenience, privacy, dependability and the like, and relative superiority of buses in avoiding parking problems and repairs, as well as ecological advantages. The 1975 mean profiles of car versus bus shown in Table 4 show that the same basic patterns of pluses and minuses were noted one year later. In general, where there was a low probability of obtaining sample means for car versus bus images due to changes for a particular attribute in 1974 (right-hand column in

TABLE 3
1974 Adults
Car Versus Bus

Variables	Car Mean	Bus Mean	F Ratio
1. Economical-Expensive	3.172	2.255	30.445 ²
2. Convenient-Inconvenient	1.655	3.359	123.587 ²
3. Brief Travel Time- Long Travel Time	.800	3.510	152.763 ²
4. Smooth Ride- Rough Ride	2.166	3.110	50.166 ²
5. Free from Weather- (door to door) Exposed to Weather (door to door)	1.966	3.386	92.962 ²
6. Easy to Socialize- Hard to Socialize	3.228	2.500	21.772 ²
7. Avoids Traffic Congestion-Gets Into Traffic Congestion	2.900	2.924	.030
8. High Status-Low Status	2.821	3.241	12.024 ²
9. Few Parking Problems- Many Parking Problems	2.890	1.614	59.738 ²
10. Flexible-Inflexible	1.786	3.386	107.594 ²
11. Uncrowded-Crowded	1.669	3.531	180.433 ²
12. Safe from Accidents- Likely to have Accidents	2.766	2.448	5.598 ¹
13. Fun to Drive-Not Fun to Drive	2.552	3.462	44.388 ²
14. Free from Repairs- Not Free from Repairs	3.303	2.062	65.326 ²
15. Safe from Dangerous People-Not Safe from Dangerous People	2.490	2.986	14.291 ²
16. High Pollution per Rider-Low Pollution per Rider	2.669	3.579	27.649 ²
17. Relaxing-Full of Tension	2.641	2.710	.252
18. Easy with Packages- Difficult w/Packages	1.731	3.648	192.355 ²
19. Can Look at Scenary- Can't Look at Scenary	2.793	1.966	42.474 ²
20. Easy to Read-Hard to Read	3.855	2.469	82.155 ²
21. Low Energy Use per Passenger-High Energy Use per Passenger	3.207	1.917	61.592 ²
22. Radio or Tape Deck Available-No Radio or Tape Deck Available	2.159	3.986	149.437 ²
23. Dependable-Undependable	1.786	2.807	58.017 ²
24. Pleasant Riding Surroundings-Unpleasant Riding Surroundings	2.207	2.945	30.893 ²
25. High Privacy-Low Privacy	1.662	4.035	321.189 ²
26. Difficult w/Children- Easy w/Children	3.531	2.628	34.735 ²
27. Quiet Ride-Noisy Ride	2.221	3.441	94.014 ²

(1=extremely, 2=moderately, 3=neutral, 4=moderately, 5=extremely)

¹p < .05

²p < .01

Table 3), there was always a low probability of attributing the perceived gap between the two modes as rated in 1975, in terms of the same attribute (right column in Table 4). In other words, where significant differences were found between the two modes' characteristics in one year, these tended to be observed in the next year.

The reason for this is implied by the relative stability

of car and bus perceptions during this time frame. For example, a discriminant analysis of the 1974 versus 1975 perceptions of car attributes for commuting indicated that there was a 67% probability that the differences obtained between profiles of mean scores were due to change. Thus, one could not conclude that car image changed significantly during this time period without taking more than an acceptable risk of being mistaken (the type-I error probability would be far more than the usual .05 level). The profiles of mean scores for car images in 1974 versus 1975 are almost identical. (The mean scores for cars and buses overtime may be deserved in different form in Tables 3 and 4. Due to lack of significant differences and space limitations tabular presentation of the discriminant analyses are not presented here.)

There was even more stability in the perceived image of buses as a commuter mode in this area. The Wilks Lambda statistic for overall discriminability of the 27 attribute ratings for 1974 versus 1975 was again nowhere near statistically significant (alpha=.85). For both the car and bus images, given the lack of overall significance between profiles for the two years, and given no more than three attributes that appeared to change significantly at the .05 level, (with 2.7 expected changes out of 54 comparisons, due to chance), it would be unwise to attempt to attribute any meaning to either of the "perceived changes" for either mode. Perhaps as conditions in the environment change more dramatically, and as more people begin to utilize public transportation, changes in the relative utility of the two major modes might be reflected in their perceived images. During the 1974 to 1975 time frame in this major southwestern city, no significant changes in perceptions can be reliably reported.

The last two tables show the only instance of real variation between the data for the two survey years, although here, again, the practical significance of these differences for transit planning purposes is quite marginal. In Table 5, we note that there were found significant changes, from 1974 to 1975, in mean desirability scores for seven of the ten evaluated financing alternatives for public transportation. All of these shifts were in the direction of lower desirability for various subsidy plans. However, the relative ordering of these alternatives was virtually unchanged from 1974 to 1975, with a Spearman rank correlation coefficient of .976 (significant at beyond the .01 level). Moreover, the frequently proposed subsidy from the "highway trust fund" is still relatively favorably received, provided the subsidy is a relatively minor part of the non-rider burden (one or two cents tax per gallon.) "No-fare" plans for riders continued to lack popular support, only "more so" than before, and attitudes toward electric bill subsidies are the most negatively perceived financing mechanism. The increasing resistance to property and electric bill tax subsidies may be partially due to increased unemployment in the study area (and in the U.S.) from 1974 to 1975, along with a dramatic increase in electric utility bills.

Table 6 suggests that the generally more conservative approach to tax and other subsidies for public transportation might also be partially due to differences in the demographic composition of the second year sample. Compared to the respondents from 1974, the 1975 group was significantly less female (50 percent versus 62 percent), older (mean age about 37 versus 35.7), longer residents in the city (mean about 6 months longer), and less educated (by about one-half year of formal education). Most of this difference is probably due to tighter controls over the representativeness of the sample, since the 1975 group is somewhat more representative of the universe. However, it should be noted that although statistically significant, most of these differences are

TABLE 4
1975 Adults
Car Versus Bus

Variables	Car Mean	Bus Mean	F Ratio
1. Economical-Expensive	3.220	2.055	41.079 ²
2. Convenient-Inconvenient	1.817	3.257	47.436 ²
3. Brief Travel Time-Long Travel Time	1.817	3.395	101.573 ²
4. Smooth Ride-Rough Ride	2.037	2.936	44.078 ²
5. Free From Weather-(door to door) Exposed to Weather (door to door)	1.826	3.321	85.681 ²
6. Easy to Socialize-Hard to Socialize	3.018	2.651	4.996 ¹
7. Avoids Traffic Congestion-Gets Into Traffic Congestion	2.945	2.807	.686 ¹
8. High Status-Low Status	2.817	3.184	8.034 ²
9. Few Parking Problems-Many Parking Problems	2.963	1.651	55.149 ²
10. Flexible-Inflexible	1.835	3.367	91.762 ²
11. Uncrowded-Crowded	1.945	3.450	110.464 ²
12. Safe from Accidents-Likely to have Accidents	2.670	2.358	5.082 ¹
13. Fun to Drive-Not Fun to Drive	2.495	3.349	29.792 ²
14. Free from Repairs-Not free from Repairs	3.184	2.110	40.438 ²
15. Safe from Dangerous People-Not Safe from Dangerous People	2.578	2.963	7.110 ²
16. High Pollution per Rider-Low Pollution per rider	2.486	3.532	34.730 ²
17. Relaxing-Full of Tension	2.514	2.606	.382
18. Easy with Packages-Difficult w/Packages	1.725	3.642	160.145 ²
19. Can Look at Scenery-Can't Look at Scenery	2.661	1.973	25.175 ²
20. Easy to Read-Hard to Read	3.569	2.385	66.397 ²
21. Low Energy Use per Passenger-High Energy Use per Passenger	3.395	1.973	63.211 ²
22. Radio or Tape Deck Available-No Radio or Tape Deck Available	2.211	3.991	109.448 ²
23. Dependable-Undependable	1.844	2.633	26.537 ²
24. Pleasant Riding Surroundings-Unpleasant Riding Surroundings	2.174	2.743	26.315 ²
25. High Privacy-Low Privacy	1.679	3.670	200.252 ²
26. Difficult w/Children-Easy w/Children	3.596	2.661	30.270 ²
27. Quiet Ride-Noisy Ride	2.211	3.119	47.352 ²

(1=extremely, 2=moderately, 3=neutral, 4=moderately, 5=extremely)
¹p < .05
²p < .01

slight, and apparently had impact more on the financing attitude profiles than on the modal choice criteria and mode images. To check this, we correlated demographic variables with the other survey questions, with particular attention given to the correlations with age, education, sex and time in the city, as large correlations with these variables might have confounded the changes(or counteracted what would have been changed, where none were reported). The results indicate that little effect can be attributed to demographic

TABLE 5
Discriminant Analysis on Financing Attitudes
1974 Versus 1975

	1974		1975		F
	Mean Attitude	Rank	Mean Attitude	Rank	ratio
Would you pay 1 or 2 cents tax/gal. of gasoline with that money going to mass transit?	2.72	1	2.81	2	.40
Riders should pay full costs of service	2.88	2	2.79	1	.40
Riders pay most costs; with balance from gasoline tax revenue	2.93	3	3.30	3	7.42 ¹
Would you be in favor of a 1/2% increase in the current sales tax with the money collected earmarked for mass transit improvement?	3.12	4	3.43	4	4.61 ¹
Would you...favor paying higher vehicle license plate fees on your personal vehicle with the money...for mass transit	3.25	5	3.52	5	3.33 ²
"No fare" for riders; mass transit financed by gasoline tax...	3.40	6	3.63	6	2.52
Riders pay most costs, with balance from tax added to property taxes	3.89	7	4.25	7	9.23 ¹
Riders pay most costs, with balance from tax on electric bills	4.03	8	4.39	9	10.64 ¹
"No fare" for riders; mass transit financed by tax added to property taxes	4.07	9	4.38	8	6.93 ¹
"No fare" for riders; mass transit financed by tax added to electric bills	4.27	10	4.47	10	3.69 ²

¹p < .05
²p < .10
 Definitely yes=1, Yes=2, Neutral=3, No=4, Definitely no=5.
 $r_s = .976$
 Probability (one-tailed test) < .01

variations between the two samples. The highest correlation between any of these four variables and criteria for modal choice showed that more educated people sought more convenience than did less educated people. Thus, convenience should have become less determinant in 1975, given a less educated sample (and it did drop slightly, but not significantly). However, the shared variation between education and need for convenience was less than 8 percent ($r = .28$), and may have been partially offset by the negative correlation between convenience and time in city ($r = -.136$), and the slightly longer average time in the city for the Year Two group. The vast majority of correlations between demographic variables (where sampling differences were found), modal choice criteria, and modal perceptions were not statistically significant, and where so, involved between 3 percent to 5 percent shared variance between demographic variations and choice criteria. Thus, for practical purposes, one

TABLE 6
Discriminant Analysis on Demographic Profiles
1974 Versus 1975

Variable	1974 Mean	1975 Mean	F-ratio
Sex (1=M,2=F)	1.619	1.503	5.401 ¹
Marital Status (1=Single, 2=Married)	1.825	1.837	.032
Student Status (1=Full time student, 2=Part time student, 3=Not student)	2.635	2.648	.031
Age (<21, 21-29,30-44, 45-59,>60)	2.814	3.044	3.764 ¹
Household Size (1=1, 2=2, 3=3, 4=4, 5=5)	2.861	2.786	.319
Education (1=Jr Hi, 2=Hi sch., 3=Hi sch. grad, 4=College/Prof., 5=College grad)	3.758	3.509	3.921 ¹
Income (<5,000, 5,000-9,999, 10,000-14,999, 15,000- 19,999, >20,000)	2.429	2.459	.055
Number of Autos (1=none, 2=1, 3=2, 4=3+)	2.571	.440	2.343
Time in Austin (<6 mo, 6 mo-1yr, 1-3yr, 3-5 yr, 5yr+)	4.179	4.434	5.264 ¹
Work Downtown (1=yes, 2=no)	1.778	1.755	.291
Shop Downtown (2/wk 2-3/mo 1/mo every 2-3mo. almost never)	3.833	3.925	.413
Shop Highland Mall (same scale as above)	3.270	3.333	.251
Shop Hancock Center (same scale as above)	3.468	3.491	.032
Shop Southwood Center (same scale as above)	4.452	4.434	.029

¹_p < .05
Number in 1974 = 252
Number in 1975 = 159

could take either 1974 or 1975 data as representative of the average resident's criteria in modal choice. (The exception would be that non-response bias affects both groups and probably overstates the receptivity to public transit and its funding, but this effect is probably constant throughout the time period.)

Inspecting the correlations between these four demographics and financing attitudes also showed some slight correlations, although again most were not significant. For example, the correlation between the willingness to pay one or two cents per gallon from the gasoline tax as a public transit subsidy (which both years' data indicate is the preferred subsidy method, if there is going to be one) correlated .038 with sex, .043 with age, -.108 with education, and .147 with time in the city. With such small correlations, little impact on the mean attitude towards this method can be attributed to demographic fluctuations in the samples. However, the less educated and longer-in-city 1975 group might have raised the mean slightly (indicating slightly less favorable attitude than before). Similarly, the

slightly stronger correlations between these demographic variables and attitudes toward substantial use of the highway trust fund versus riders paying the entire cost, shows a slightly greater effect of more conservative demographics in 1975 influencing some of the slight shift in this direction regarding financing alternatives. However, again, the highest demographic correlate is .248, or about 6 percent shared variation between "time in city" and resistance to using the highway trust fund for public transit. The overall attitude toward this mechanism was still neutral to positive, and whatever impact these small correlations with demographic variations might have had on some mean score shifts between 1974 and 1975 was minimal. The relative preferences for financing alternatives were virtually unchanged, and modal choice criteria and perceived mode images were similarly stable.

Conclusions

In spite of fluctuations in the general perceived importances of the inflation, the economy, and "energy crisis" from Spring, 1974 to Spring, 1975 (Opinion Research Corporation, 1974; 1975), there seems to have been little movement in the basic determinant attributes of modal choice. Further, the general image gaps between private versus public transportation modes in this study area remained roughly constant over the interval. Given the rather large disparity between cars versus buses for the average respondent, who has considerable discretion, it is not surprising that significant modal switching did not occur. Our data indicated a modest increase in car-pooling activities and slightly bigger car purchases. Given the current environmental conditions, and perceived benefits of private versus public transportation (especially in dependability, convenience, and flexibility) for this survey area, relatively major changes will be needed either in the perceived attributes of public transportation (or perhaps over time in the relative determinance of attributes in which public transportation is already seen as superior) for any amount of modal switching to occur. At the margins some potential switchers may be converted by treating them as a distinct market segment. The general procedure would involve those persons now using private transportation who indicate a high likelihood that they would switch to an unproved public transportation system. An approach describing a procedure for the identification of potential switches and their determinants of modal choice, plus appropriate promotional strategies aimed through media to which these persons were shown to be differentially exposed, is discussed in Alpert and Davies (1975).

Aside from this relatively small group as likely switchers, the average respondent is still rather far from altering his/her life-style to the extent that conversion to public transportation implies at present. Implementing changes in the attributes shown as determinants of modal choice in both years is more likely to bring about shifts in travel patterns and modal choices. Many of these changes may be expensive, but our data suggest a public willingness for tapping the highway trust fund, particularly if increased utility in public transportation is produced, in terms of attributes sought by the travelers (and to some extent lacking in private transportation).

If further research indicates the same kind of stability in modal choice criteria as was shown here, improvements in the system characteristics will become even more important for generating behavioral changes. If public desires are stable, the system must become more responsive; the data presented here indicate that criteria are not changing in any significant way so far. This simplifies the planning process, although it may also suggest promotional campaigns to alter criteria. At present, the available public transportation system is

not seen as sufficiently competitive to private alternatives. In this study area, the gaps were not closed from 1974 to 1975. Additional research in other areas should be done to check for generality. In addition, it is possible that trends may be observed in further research that can monitor the public's modal choice criteria and their evaluation of alternative modes' abilities to meet them, as both programmed and unplanned changes occur in the relevant transportation system and its environment.

References

- Mark I. Alpert, "Identification of Determinant Attributes: A Comparison of Methods," Journal of Marketing Research, 8 (May 1971), 184-91.
- Mark I. Alpert and Shane Davies, "A Decision Sciences Approach to the Marketing of Public Transportation," presented at AIDS Sixth Annual Meetings (November, 1974).
- Robert C. Blattberg and Samuel R. Stivers "A Statistical Evaluation of Transit Promotion," Journal of Marketing Research, 7 (August 1970), 293-99.
- "A New Federal Stab at Aiding Mass Transit," Business Week, January 16, 1974, 52-53.
- Thomas Golob, Ricardo Dobson and Jagdish N. Sheth, "Perceived Attribute Importance in Public and Private Transportation," American Institute for Decision Sciences Proceedings, 5 (November 1973), 7-10.
- Phillip Kotler, Marketing Management, 3rd ed. (Englewood Cliffs, N.J.: Prentice-Hall, 1976).
- D.M. Hill and H.G. Von Cube "Development of a Model for Forecasting Travel Mode Choice in Urban Areas," Highway Research Record 38 (1963), 78-96.
- Stanely J. Hille, F.T. Paine, A.N. Nash and G.A. Brunner "Consumer Transportation Attitude in Baltimore and Philadelphia," Transportation Journal, (Summer, 1968), 30-47.
- Stanley J. Hille, and T.K. Martin "Consumer Preferences in Transportation," Highway Research Board Record (1967), 6-43.
- R.A. Mundy, D.W. Cravens and R.B. Woodruff "Potential for Marketing Management Applications in Public Transportation Planning," American Marketing Association Proceedings (1974), 274-279.
- James H. Myers and Mark I. Alpert "Determinant Buying Attitudes: Meaning and Measurement," Journal of Marketing 32 (October, 1968), 13-20.
- James H. Myers and W. Gregory Warner, "Semantic Properties of Selected Evaluation Adjectives," Journal of Marketing Research, 4 (November, 1968), 426-428.
- Opinion Research Corporation (Gallop Poll), 1974, 1975.
- Jagdish N. Sheth, "A Psychological Model of Travel Mode Selection," Association for Consumer Research Proceedings, (1975), 425-30.
- Donald J. Veldman, Fortran Programming for the Behavioral Sciences, New York: Holt, Rinehart and Winston, (1967).

ENERGY CONSERVATION, PRICE INCREASES AND PAYBACK PERIODS*

William H. Cunningham, The University of Texas at Austin
Brondel Joseph, The University of Illinois at Chicago Circle

Abstract

This research examines consumers' responses to price increases in gasoline and electricity as well as their willingness to accept various payback periods for investments in insulation and solar energy equipment. The results indicate that consumers will conserve energy in response to significant price increases in energy although they are not willing to wait extended periods of time to recover investments in energy saving equipment.

Introduction

Examination of consumer response to various incentives designed to encourage energy conservation is just beginning. Preliminary findings in the area are not encouraging. The response given first and foremost by consumers is that the most acceptable energy policies are those which require the least inconvenience, the least personal cost, and the least change in lifestyle (Cunningham and Lopreato, 1977; Doering *et al.*, 1974; Gottlieb and Matre, 1976; Grier, 1976). Bartell (1974) reports greatest public acceptance for behavioral regulation such as the 55 m.p.h. speed limit, a required reduction of 10 percent in electricity use, and reserved freeway lanes for buses and car pools. In addition, several studies are available which deal with the impact of price adjustments on various income sectors of society (Berman and Graubard, 1973; Berman, Hammer, and Tihansky, 1972). The major suggestions emerging from these analyses are, not surprisingly, that high-income groups display a greater ability than low-income groups to reduce electrical consumption, both direct and indirect energy expenditures are regressive in nature, and the ability of low-income groups to reduce their consumption of energy is relatively low.

In summary, not very much is known about consumer reactions to conservation incentives. The objectives of this research are first to examine consumers' reactions to price increases in gasoline and electricity and second to analyze the payback periods that consumers are willing to accept to invest in insulation and solar energy. The data is broken down by income groups because federal legislation has been proposed with energy incentive provisions which impact on family's taxable income. Also most incentive systems are in some way based on price and therefore it seems reasonable to focus the present research on consumers' ability to respond to the price mechanism.

Methodology

The data for the study were collected through a mail questionnaire sent to residents in five communities: Austin and El Paso, Texas; Flagstaff and Prescott, Arizona; and Albuquerque, New Mexico. These cities were selected to represent various types of communities in the Southwest. The subjects were selected from the consumer billing records of the local natural gas company. Every tenth residential consumer on the firm's mailing list was chosen for the sample. Twelve percent of the residential customers serviced by the electric utility in Flagstaff and Prescott are all-electric home customers; three percent of the customers

of the utility servicing Albuquerque are all-electric customers, as are nine percent in Austin, and one percent in El Paso (Federal Power Commission, 1975). While it cannot be assumed that all of the remaining residents supplement their electrical energy with natural gas, it is highly likely in this region of the country that the vast majority use some natural gas and are therefore part of the population sampled.

Ten thousand questionnaires were mailed with an accompanying letter explaining that consumers were being asked to participate in a study sponsored by The University of Texas at Austin and that responses would be reported only in aggregate form. A response rate of nearly 25 percent yielded 2,403 codable returns for analysis.

The subjects were broken down into the following six groups based on their total annual family income: (1) <\$5,000; (2) \$5,000-9,999; (3) \$10,000-14,999; (4) \$15,000-19,999; (5) \$20,000-24,999; (6) >\$25,000. They were asked to state what their reactions in terms of fuel use would be to nine price increases varying from five percent to more than 150 percent in gasoline, and electricity. The response categories which were presented included: (1) "no reduction," (2) "slight reduction," (3) "moderate reduction," (4) "substantial reduction," (5) "maximum possible reduction," and (6) "would no longer use." The data are presented such that a score of 1 equals no reduction, a score of 2 equals a slight reduction, and so forth.

The subjects were also questioned on what would be the maximum amount of time that was acceptable to them to recover various levels of investments in insulation and solar energy equipment through savings in their energy bills. The response categories included: "less than one year," "1-2 years," "3-4 years," "5-6 years," "7-8 years" and "more than 8 years." F-tests were used to determine if significant differences existed between the respondent groups, with respect to their reactions to each price increase and payback periods.

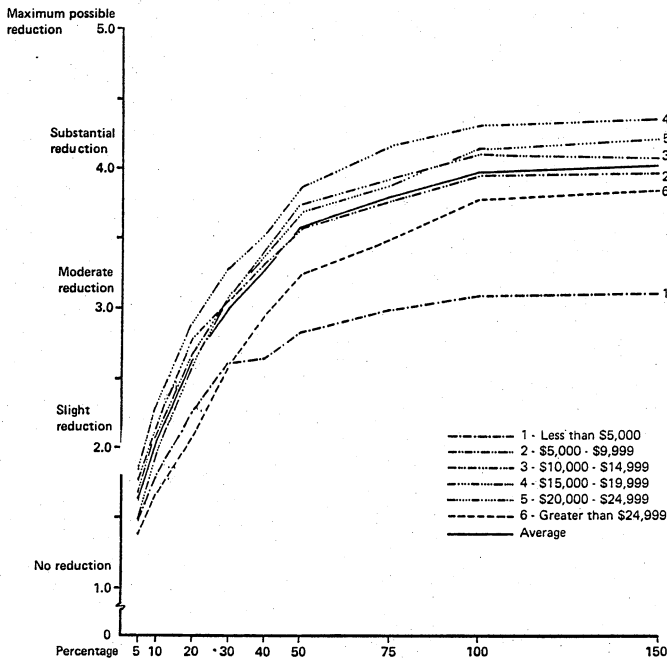
Results

Price Increases in Gasoline

Figure 1 compares the mean responses of the total sample along with the mean responses of each of the six income groups to price increases in gasoline ranging from 5 percent to 150 percent. It is apparent that even a 5 percent increase in the price of gasoline elicits some energy-conserving reactions on the part of all income groups. A 10 percent increase generates a mean response from all consumers of 1.95, just below the "slight reduction in the use of gasoline" category. A 30 percent increase in the price of gasoline elicits a mean response score of 3.01, which is virtually equivalent to "moderate reduction in the use of gasoline." A 50 percent increase elicits an average response of 3.58, which is midway between moderate and substantial reduction. It is important to note that while there is wide variation by income group, the average score for the total sample on a 100 percent price increase is 4.0, which is equivalent to "substantial reduction in the use of gasoline."

Figure 1

Consumers' Responses to Percentage Price Increases in Gasoline



The patterns reflected in Figure 1 are interesting, particularly since they are basically consistent with the data on electricity. In the first place, a 50 percent price increase marks a threshold level in terms of consumer response. Price increases up to 50 percent elicit markedly rising levels of reported conservation. After the 50 percent mark, however, the response levels off, and little further conservation is reported. It is not possible to determine whether this pattern is due to consumers' refusals to cut back on gasoline use past "substantial" reduction, or whether consumers simply find it difficult to believe that price increases up to 150 percent over present levels would occur. Average gasoline prices in the southwest had already risen 30 percent from October 1974 to October 1975 when the survey was taken. Consumers may have viewed the hypothetical increases above 50 percent as unrealistic.

A second pattern involves the low price responsiveness of the lowest and the highest income groups, or conversely, the high responsiveness of the middle-income consumers. This finding is consistent with earlier work (Grier, 1976). The assumption is widespread that low-income groups cannot do much about their energy use and that high-income groups will not. The present findings tend to bear out that assumption.

These consumers whose total family income is below \$5,000 per year are by far the least price responsive. Their conservation efforts tend to level off after the 30 percent price increase in gasoline. The gap between the responses of this group and those of the other groups is substantial. The difference between the mean response score of the lowest income group and the mean score of the sample at the 100 percent hypothetical price increase, for instance, is .91 (4.0-3.09). It is difficult to say why the pattern of response is not linear. That is, why is it that low-income people do not indicate "maximum" reduction in use at 30 percent

price increases instead of responding with only "moderate" reduction? Any explanation here is post hoc, but it may reflect the attitude that "there is nothing more we can do." Low-income people use much less fuel than other groups of consumers, so that a moderate reduction in that use does not signify the same thing as a moderate reduction among more affluent consumers.

Consumers with yearly family incomes above \$25,000 are the least price responsive at the lower levels of increase (5-30%). At the 30 percent level they become more responsive than the low-income group, but it is not until the 100 percent increase that they draw close to the middle-income groups. At all hypothetical price levels, the consumers in the \$15,000-19,999 category show the most reported behavioral change. The differences in response by income group are statistically significant (see Table 1) as well as substantively clear from the graphs. It is apparent that price increases are much more effective incentives for some groups of people than for others, and that the high energy users are not those most affected.

Table 1

Results of Analysis of Variance for Gasoline and Electricity Price Increases

% Price Increase	F-ratio ^a Gasoline	F-ratio ^a Electricity
5%	7.34	6.28
10%	10.49	5.48
20%	11.41	7.01
30%	9.16	6.33
40%	9.74	8.37
50%	11.57	9.35
75%	12.19	9.88
100%	11.60	11.50
150%	10.45	11.41

^aAll F-ratios are significant at the $p < .000$ and all have 5 degrees of freedom.

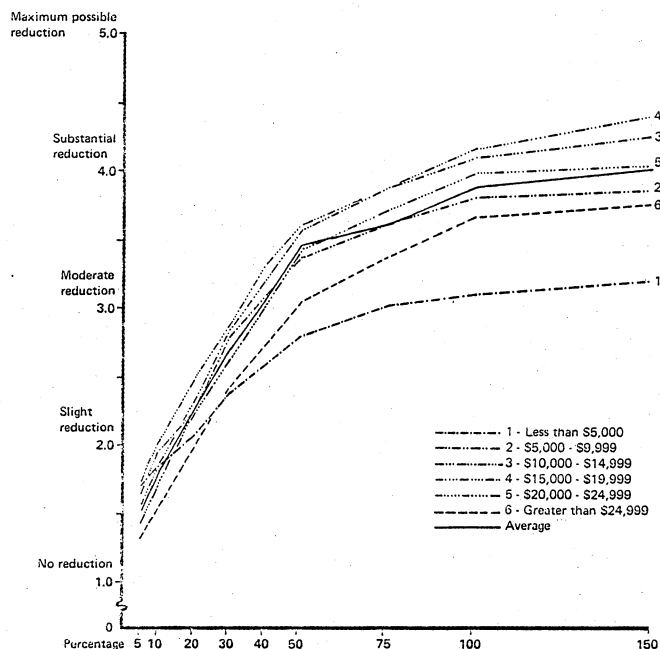
Price Increases in Electricity

The responses of consumers to price increases in electricity are quite similar to their responses for price increases in gasoline. Figure 2 indicates that the mean response for the total sample for percentage price increases in electricity is practically linear for 5 percent to 50 percent. Consumers are willing to make only the most modest conservation adjustment to a 5 percent price increase. However, they appear increasingly price responsive up to a 50 percent price increase, which elicits a mean response for the entire sample of 3.4, substantially in excess of "modest reduction in the use of electricity." Increases above 50 percent elicit additional conservation, although the behavior curves begin to level off. For the 100 percent to 150 percent price increases, the sample's mean response barely changes.

Analysis of variance significantly differentiates the income groups based on responses to the various price increases (see Table 1). The lowest income group (<\$5,000) is the most responsive to a 5 percent increase in the price of electricity, with a mean score of 1.72. This pattern quickly shifts, however, with additional price increases in the range from 30 percent to 150 percent.

Figure 2

Consumers' Responses to Percentage Price Increases in Electricity



With two exceptions (the 5 percent and 75 percent price increases), the most responsive subjects to price increase were in the \$15,000 to \$19,999 income group. For both 100 percent and 150 percent price increases, the consumers in this group indicated, as a whole, that they would make more than a "substantial reduction in the use of electricity."

The patterns of response to price increases in electricity are similar to those for gasoline. The highest and lowest income groups are the most responsive. Degree of reduction in use is virtually the same for the two fuels.

Payback Periods for Insulation

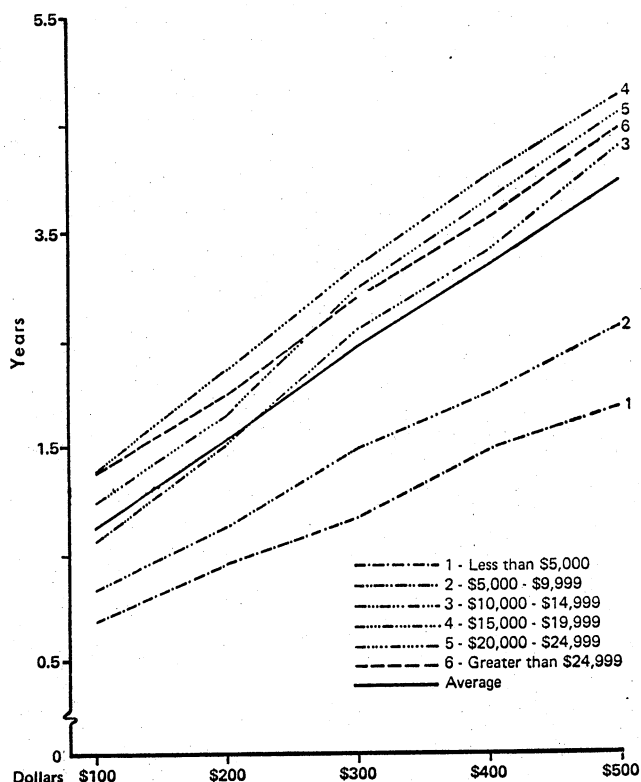
The subjects were presented with five investment capital alternatives for insulation: \$100, \$200, \$300, \$400, and \$500. As Figure 3 indicates, the mean responses for the entire sample show a linear pattern. For an investment of \$100 the sample as a whole indicated that investment should be recovered in approximately one year. Each \$100 of additional investment means that investors will wait approximately eight additional months for recovery. As a result, for an investment of \$500 the sample, as a whole, responds that investment recovery must be made in approximately four years.

It is interesting to note from Figure 3 that the six income groups are dichotomized in their responses. On one hand there are those consumers in the two lowest income segments (<\$5,000 and \$5,000-\$9,000), who clearly must be able to recover even small amounts of investments in insulation very quickly. As an example, the lowest income segment feels it can wait no longer than approximately six months to recover expenditure of \$100 on home insulation. For an investment of \$500 this group would be willing to wait only a little longer than one and one-half years to recover the investment.

On the other hand, the four wealthiest segments of the population are rather similar in their behavior towards

Figure 3

Maximum Acceptable Time to Recover Investments in Insulation



the insulation. They are willing to accept longer periods of time to recover investments from savings in insulation. As an example, the fourth highest income segment (\$15,000-\$19,000) is consistently willing to wait longer than any other income group to recover the cost of insulation. For an investment of \$100 they feel they need to recover this investment in just under one and one-half years, while an investment of \$500 finds them willing to wait more than four and one-half years to recover their expenditures on insulation. Table 2 indicates that the six income groups are significantly different in terms of their responses.

Table 2

Results of Analysis of Variance for Capital Investments in Insulation and Solar Energy

Investments Alternatives	F-ratio ^a Insulation	F-ratio ^a Solar Energy
\$100.00	13.05	---
\$200.00	15.85	---
\$300.00	19.23	---
\$400.00	18.95	---
\$500.00	20.57	7.32
\$1,000.00	---	13.99
\$3,000.00	---	15.50
\$10,000.00	---	22.23
\$15,000.00	---	23.57

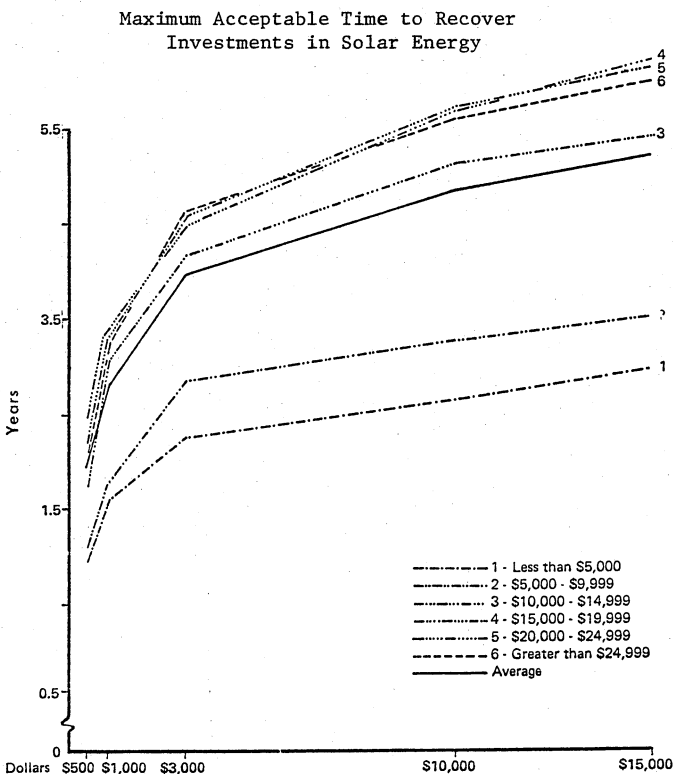
^aAll F-ratios are significant at the p < .000 level and all have 5 degrees of freedom.

Payback Periods for Solar Energy

The subjects were asked what would be the maximum amount of time that they would be willing to wait to recover capital investment through savings in energy bills from solar energy equipment. The required investments were: \$500, \$1,000, \$10,000, and \$15,000.

As Figure 4 indicates, the mean responses to the solar energy investments are curvilinear. For a \$500 investment the subjects are willing to wait slightly more than one and one-half years. In the same manner, for a \$1,000 investment the average acceptable time period is more than two and one-half years. However, as the investment increases in size the consumers are not willing to wait for proportional time period increases to recover their investment. For example, a \$10,000 investment finds the average subject willing to wait only slightly more than four and one-half years. For a \$15,000 investment the average subject is willing to wait only a little longer than five years.

Figure 4



Once again, the lowest two income segments tend to react in the same manner while the highest four income segments also tend to respond to the questions similarly. Although this pattern of responses also occurred in the examination of the insulation data, it is much more pronounced with the solar energy data.

The responses for the two lowest income groups change substantially from \$500 to \$1,000 to \$3,000 investments. As an example, the second highest income group felt it should recover its \$500 investment in solar energy in slightly under one and one-half years while for an investment of \$3,000 they are willing to wait more than two and one-half years. However, for investments from \$3,000 to \$15,000, the lowest two income segments do not vary greatly. The second lowest group is willing to wait only three and one-half years to recover a \$15,000 investment in solar energy.

In contrast with the pattern of the two lowest income segments, the upper-income segments were more willing to accept substantially longer time periods to recover their investments in solar energy equipment. As Figure 4 indicates, there are few differences among the response patterns for the three highest income groups. The third highest income segment's behavior is very similar to the top three, although it is distinguishable in all cases. The highest three income groups are willing to wait more than four and one-half years to recover a \$3,000 investment in solar energy equipment. For an investment of \$15,000, each of the three highest income segments are willing to wait more than five and one-half years. Table 2 indicates that there were significant differences among the income groups' responses to the solar heating and cooling investment questions.

Implications

There are several important implications from the present study. First, conservation of gasoline and electricity can be achieved with price increases. While this finding is not surprising, it is important to know the price levels which are most encouraging of conservation. In each case an increase of 50 percent appears to be the point below which substantially less conservation occurs and above which little additional conservation occurs. It is, however, interesting to note that President Carter's proposed five cent per gallon gasoline tax, which would be imposed in January of each year when the nation's consumption of gasoline exceeds the previous year's target, seems to result in little more than a redistribution of income. The data is quite clear that a five cent price increase per gallon of gasoline will not have a material effect on gasoline consumption.

Second, it is possible to identify different conservation patterns on the part of the six income groups to changes in the price of the gasoline and electricity. Although each income group displayed a curvilinear reaction to the price increases, the \$10,000-\$14,999, \$15,000-\$19,999, and \$20,000-\$24,999 income groups were generally the most willing to conserve energy. As an example, at price increases from 75 percent to 150 percent these groups scored above the average for the entire sample on conservation efforts for both fuels. In contrast, the lowest income segment (<\$5,000) was normally the least willing to conserve; the next least willing to conserve was the wealthiest segment of society. Although it is not possible to know exactly why this pattern takes place, it is reasonable to speculate that the low-income consumers are already conserving energy and are just not in a position to conserve more energy, while the upper-income people are simply not as concerned about the cost of energy as is the sample as a whole.

Third, as was the case for responses to price increases, clear patterns emerge in responses concerning recovery time for costs of energy-saving investments. In this case, however, the two lowest income groups are distinctly split off from the other groups. Low-income consumers do not seem likely to make investments that require long payback periods. The middle and upper-income groups are more willing to wait longer periods of time to recover their investment. However, even for large investments (on solar energy equipment, for example), consumers generally expect a recovery time of less than six years. This seems to indicate that long-term life cycle energy savings from home improvements will not be effective inducements and that more direct approaches must be made to lower initial investments.

Conclusions

The present study has shown that consumers seem to be willing to conserve energy in response to price increases. This may appear to be in conflict with the recent study done by Willenborg and Pitts (1977) which found that the price mechanism was not very effective in reducing consumption of gasoline when prices are increased gradually over time. However, the present study was done after the price increases that Willenborg and Pitts were studying had taken place. In addition, the present study does not imply that the mandated price increases must take place gradually. This research does seem to indicate that the price mechanism will work quite effectively up to a 50% price increase. After this point, if further conservation efforts are required, additional steps such as allocation of fuels will have to be employed.

The study also found that consumers are not willing to wait very long to recover their investments in energy saving equipment. Since the retrofitting of homes will probably not result in payback schedules which are acceptable to most consumers the federal government should consider the use of its tax depreciation, direct tax write off, and income credit powers to stimulate the adoption of energy saving equipment.

References

- Ted Bartell, "The Effects of the Energy Crisis on Attitudes and Lifestyles of Los Angeles Residents," Proceedings (American Sociological Association, Summer Conference, 1974), pp. 88-93.
- M. B. Berman and Morlie H. Graubard, "A Model of Residential Electricity Consumption," Rand Paper Series, (P-5063) (July 1973).
- M. B. Berman, M. J. Hammer and D. P. Tihansky, The Impact of Electricity Price Increases on Income Groups: Western United States and California (Santa Monica: Rand Corporation (R-1050-NSF/CSA), 1972).
- William H. Cunningham and Sally Lopreato, Consumers Energy Attitudes and Behavior, (New York: Praeger Publishers, 1977).
- Otto Doering, Jerry Fezi, Dave Gauker, Mike Michand and Steve Pell, Indiana's Views on the Energy Crisis (Lafayette: Purdue University Agricultural Economics Department, Cooperative Extension Service, 1974).
- Federal Power Commission, All Electric Homes in the United States (Washington: United States Printing Office, 1975).
- David Gottlieb and Mari Matre, Sociological Dimensions of the Energy Crisis--A Follow-up Study (Houston: The University of Houston, The Energy Institute, 1976).
- Eunice Grier, "Changing Patterns of Energy Consumption and Costs in U.S. Households," (Allied Social Science Associations Meeting, September 1976).
- John F. Willenborg and Robert E. Pitts, "Gasoline Prices: Their Effect on Consumer Behavior and Attitudes," Journal of Marketing, 41 (January, 1977), pp. 24-31.

*Adopted from Energy Use and Conservation Incentive, with permission of Praeger Publishers, New York, 1977.

AN EXPERIMENTAL EXAMINATION OF DECEPTION IN LABELING:
CONSUMER RESEARCH AND PUBLIC POLICYMAKING

Philip G. Kuehl, University of Maryland
Robert F. Dyer, George Washington University

Abstract

This study expands knowledge about potential sources of marketplace deception in two ways. First, conceptual and methodological insights evolving from research in deceptive advertising are applied to the empirical study of deception in labeling. Second, since existing U.S. Department of Agriculture labeling regulations affecting actual consumer meat products are examined in the study, consumer research input to public policymaking processes are illustrated.

Introduction

The subject of deception in advertising, packaging, and labeling has emerged as a highly visible public policy-related issue among consumer researchers in recent years. Three major factors have contributed to such widespread interest in examining marketplace practices which exhibit the potential to deceive consumers.

First, the so-called consumer movement has focused much of its recent emphasis on encouraging business firms and public sector institutions to implement policies (i.e., the "right to be informed") which strengthen the availability of consumer information in the marketplace. Second, public sector institutions like the Federal Trade Commission (FTC) and Food and Drug Administration (FDA) have pursued policies to reduce deceptive practices in the marketplace. Similarly, agencies like FDA and the U.S. Department of Agriculture (USDA) have funded large-scale consumer research studies to examine a variety of product packaging and labeling policy alternatives (e.g., open-dating provisions; product storage and usage instructions; disclosure of ingredients statements; nutritional labeling; "package inserts" with consumer information; and "safety" caps and injury prevention packaging in poisons, medicines, etc.). Third, many companies have initiated activities in support of consumers' rights to product information in the marketplace by sponsoring innovative consumer education programs. In the Washington, D.C. metropolitan area, for example, Giant Foods, Inc. has embarked on a series of such consumer education activities (under the leadership of Esther Peterson who is now Special Assistant to the President for Consumer Affairs) which focus on improving the consumers' ability to make intelligent purchase decisions.

Objectives

While the preceding factors highlight general concern about alleged deceptive practices in the marketplace, conceptual and empirical studies published by consumer researchers to date have focused primarily on the issue of deception in advertising. As a result, there appears to be a need for research which expands this existing research tradition to include other sources of potential deception. The present study helps to fulfill this need by providing insights on two major objectives.

1. To determine if conceptual and methodological knowledge about deception in advertising is applicable to measuring deception in consumer product labeling

2. To provide consumer research input to public policymaking processes by examining whether or not existing USDA-approved labeling is likely to stimulate falsely based consumer beliefs associated with the attributes of two alternative meat products

Since the conceptual objective of the study relates to over 25 major articles dealing with defining deception in advertising, the following discussion summarizes this body of knowledge and assesses its applicability to measuring deception in labeling. Next, the public policy related objective of the study--which involves actual USDA labeling regulations for "ham" and "turkey-ham" products--is discussed. The study methodology, findings, and conclusions are presented in the remaining three major sections of the paper.

Conceptual and Methodological Framework

The framework for the conceptual and methodological approaches used in this study evolves from published literature on deceptive advertising. This academic research tradition can be classified into three major types of contributions.

Preliminary Methodological and Conceptual Insights

Results from four studies in the literature (Hunt, 1972 and 1973; Dyer and Kuehl, 1974; and Kassarian, Carlson, and Rosin, 1975) provided two basic insights on conceptualizing and measuring deception from a methodological perspective. First, all of these authors recognized that experimental research designs (*vis-a-vis* cross-sectional studies) are the most appropriate methodological framework for investigating deception. In other words, these studies illustrate a need for empirical research in consumer deception to be based upon "cause and effect" research designs--where degrees of alleged deception can be directly related to specific stimuli (i.e., either a product label or an advertisement).

The second methodological insight from these early studies in deceptive advertising concerns the selection of appropriate dependent variables to measure deception. In all of these studies, such broad or "global" variables as consumer attitudes toward the "product brand" or "brand manufacturer" were used. In other words, traditional aggregate consumer response variables were used in these studies. While this measurement approach reflected the "state of the art" at that time, other academicians recognized a need to conceptualize innovative variables for examining deception which are more precise and reliable in nature.

Defining Deception in Conceptual Terms

Conceptual contributions to defining consumer deception encompass two major themes. First, many authors (Aaker, 1974; Armstrong and McLennan, 1973; Cohen, 1969 and 1972; Dillon, 1973; Permut, 1974; Ford, Kuehl and Reksten, 1975; Rosch, 1975) developed typologies for classifying and measuring characteristics of alleged deceptive practices. In essence, these authors have illustrated that the term "deception" is multidimensional in nature and that it is possible to define a variety of different forms of deception (i.e., the "blatant lie,"

the "partial truth," "claim-fact discrepancy," the "false certification," etc.). These discussions suggest that empirical research on deception must recognize the specific type or form of deception represented in the alleged misleading stimuli *vis-a-vis* adopting an aggregate meaning for the term "deception."

Second, several other authors (Jones, 1971; Wilkie and Gardner, 1974; Jacoby and Small, 1975; Brandt and Preston, 1976) developed conceptual frameworks for examining deception within the policymaking processes of public sector institutions like the FTC and FDA. In essence, these authors developed approaches for empirically defining consumer deception in an operational context that is consistent with traditional legal doctrine based approaches.

Brand Beliefs: Concepts and Research

The preceding two dimensions of academic work on defining and measuring consumer deception were valuable in stimulating a growing professional dialogue in the subject area. However, it was not until the contributions of Wilkie and Gardner in 1973 that the "belief" component of the generalized Fishbein attitude model was recognized as an appropriate and precise variable to measure the effects of alleged deceptive practices in the marketplace. In proposing the use of brand attribute "beliefs" *vis-a-vis* overall attitudinal variables to study deception, these two authors suggest that consumer beliefs about specific product attribute claims in specific marketing stimuli should be used to measure deception. Related to this study, a product label may be viewed as stimulating an "incorrect" or "misleading" consumer perception about selected product attributes (e.g., the ingredients of the product) but may not mislead the consumer on other product attributes (e.g., the nutritional value of the product).

The conceptual efforts of Wilkie and Gardner have spawned four research efforts to date (Dyer and Kuehl, 1974; Mazis and Adkinson, 1976; Kuehl and Dyer, 1976; and Kuehl and Dyer, 1977). In each of these studies, classic experimental research designs have been used to measure consumer "beliefs" about specific claims made about the product in a firm's promotional efforts. Further, the relatively consistent results from these studies support the views of Wilkie and Gardner that consumer beliefs--measured within an experimental research design--should be used to assess the impact of alleged deceptive marketing stimuli.

Summary Conclusions

The preceding discussion has highlighted the major conceptual and empirical contributions which consumer research has made to the study of deception in advertising. In relating this work to the present study, the following summary conclusions are appropriate:

1. An experimental research design should be used to study deception--so that consumer response measures can be directly related to specific claims made in product labeling through a "cause and effect" methodology.
2. Consumer "beliefs"--examined on a specific product attribute basis--should be used to measure deception in labeling.
3. The specific type of deception reflected in alleged misleading statements about the product must be identified. In this study, it appears *a priori* that labeling used for "turkey-ham" meat products is a "claim-fact" discrepancy. In other words, consumer perceptions of the "claim" that "turkey-ham" is made from a combination of meat and poultry products are not supported by the

technical "fact" (i.e., this product is made only from turkey meat).

4. The research should be consistent with viable public policymaking alternatives in the agency with jurisdictional responsibility for the alleged deception. In this study, for instance, deception in labeling should be related to potential, (a) pre-market, (b) post-market, and (c) consumer education policy actions at USDA.

Public Policy Framework

The framework for the public policy objective of the study encompasses two classes of issues. From a general perspective, it should be noted that product labeling--as a point-of-sale marketing stimulus--is receiving increased attention among public policymakers at the present time. For instance, FTC recently has instituted hearings on "information disclosure" proposals affecting non-prescription drug product labeling regulations. Similarly, the FDA had undertaken a systematic review of food product labeling policies related to provisions of the Fair Packaging and Labeling Act. As a result, research which examines product labeling as a potential source of consumer deception contributes knowledge in support of this increased policymaking activity.

From the perspective of USDA policymaking, the study extends the work of Miller, Topel, and Rust (1976). Their empirical work, which focused on USDA labeling regulations affecting beef grading, supported the view that there is a need to investigate the meat labeling practices and that USDA should reconsider the "... role it plays as a source of information to food marketers and consumers alike" (Miller, Topel, and Rust, 1976; p. 31).

In relating this viewpoint to the present study, the USDA's promulgation of two regulations affecting the labeling of "ham" and "turkey-ham" meat products provide the specific focus for the experiment. First, USDA policymakers have adopted a dictionary-based definition for the technical term "ham" (i.e., the thigh cut of meat from the hind leg of any animal). Second, USDA has approved the terminology "turkey-ham" for the labeling of a meat product which includes only turkey meat (i.e., no "ham" meat).

The justification for this label wording approval, according to USDA, is based on two major policy arguments. First, since the term "ham," in USDA parlance, refers to a cut of meat (i.e., the thigh portion of the hind leg of any animal) rather than a type of meat (i.e., smoked and cured pork meat), it is technically correct to associate "turkey" and "ham" wording on a product label. Second, while turkeys have only two legs, USDA has decided that the wings of this bird constitute two additional locomotive extremities on the body. As a result, meat taken from the thigh portion of the turkey leg can be labeled with the term "ham."

Although this USDA logic appears unduly naive from a consumer research view, the major issues in USDA policymaking are clarified by the following questions:

1. Do consumers, as a result of their exposure to USDA approved product labeling, exhibit beliefs about specific "turkey-ham" product attributes which are not supported by factual evidence on this product?
2. Do consumers, as a result of their exposure to USDA approved product labeling, exhibit beliefs about specific "turkey-ham" product attributes which are supported by factual evidence on this product?

3. What are consumer perceptions or understanding of the USDA definition of the term "ham"?

Methodology

The Experimental Design

The basic experimental design used in the study is an "after-only control group" approach. This design was chosen for three major reasons. First, this design contains two basic factors needed for classic experimental research (i.e., (1) "control" and "experimental" groups, and (2) random assignment of respondents to the two groups). Second, this design is particularly well-suited for multi-site field interviewing since the administrative and supervisory procedures needed to support this design are not complex (i.e., as a result, potential errors in the implementation of the experiment are reduced). Third, this design is especially appropriate to examine the major policy issues in the study. In other words, "ham" product labeling can be viewed as a "control" variable in the experiment since it is the traditional meat product for which "turkey-ham" is a substitute. Furthermore, if consumers are deceived by the "turkey-ham" label wording (i.e., the experimental variable) it is likely that the effect of such deception most likely will be a decreasing market share for "ham" meat products.

FIGURE 1
The "After-Only with Control Group" Experiment Design

Pre-Experimental Data Collection (Phase 1)	Experimental Data Collection (Phase 2)	Post-Experimental Data Collection (Phase 3)
S_1	$\left\{ \begin{array}{lll} R_H & X_1 & O_1 \\ R_{T-H} & X_2 & O_2 \end{array} \right\}$	S_2

where:

- S_1 = general attitudinal data on food knowledge for products other than "ham" or "turkey-ham,"
- R_H } random assignment of all study subjects at
 R_{T-H} } each site to either the control group (i.e., "ham" package) or the experimental group (i.e., "turkey-ham" package),
- X_1 = "ham" product label
 X_2 = "turkey-ham" product label } exposure
- O_1 } = specific label attribute beliefs for the
 O_2 } "ham" and "turkey-ham" labels, and
- S_2 = sociodemographic data for all study subjects.

Study Procedures

Four interviewing sites (i.e., New York, Chicago, Kansas City, and Los Angeles) were used to implement the study design. At each site, the following procedures were used:

1. Professional interviewers were recruited and trained in the methods and procedures to be used in the study. All of the interviewers were employed by local, independent interviewing services in each city.

2. A nonprobability sample of 100 respondents was selected for the study at each site (N=400). None of the respondents were paid to participate in the study.
3. Each respondent at each site was assigned a number from "1" to "100" and the Phase 1 data (general food knowledge) was collected from all respondents.
4. In Phase 2, odd numbered respondents (i.e., the "ham" product group) examined the ham package and label. The even numbered respondents (i.e., the "turkey-ham" group) followed the same procedure by examining a "turkey-ham" package and label. After examining their respective package and label, a seven-point, seven item attribute scale was used to obtain belief data. The attributes examined in the study included:
 - a. Nutrition Beliefs
 - b. Packaging Safety Beliefs
 - c. Product Content Beliefs: "Pork" Meat Only
 - d. Product Content Beliefs: More Than One Meat
 - e. Product Content Beliefs: Some "Turkey" Meat
 - f. Product Content Beliefs: "Turkey and Ham" Meat
 - g. Product Content Beliefs: Turkey Meat Only
5. Actual product packages and labels were used in the experiment: ITT Gwaltney Ham and Rich's Turkey-Ham.
6. Phase 3 (sociodemographic data) was collected from all respondents.

Data Analysis

Traditional data analyses were conducted on the study data in the following ways:

1. Phase 1 data were: (a) tabulated to provide general insights on respondent food knowledge, and (b) statistically analyzed to ensure that the "turkey-ham" and "ham" groups were equivalent in terms of general food knowledge.
2. Phase 2 data were analyzed by T-Tests which statistically examined whether or not consumer beliefs about "turkey-ham" and "ham" are significantly different as a result of exposure to the two product labels.
3. Phase 3 data were: (a) tabulated to provide a general profile on overall study respondent sociodemographic characteristics, and (b) statistically analyzed to ensure that the "turkey-ham" and "ham" groups were equivalent in terms of sociodemographic characteristics.

Results

Table 1 presents the product attribute beliefs of respondents which were generated through the experimental procedures used in the study. This discussion is followed by findings illustrating the general food knowledge and sociodemographic characteristics of the respondents. The discussion is organized in terms of the major public policy issues presented previously.

TABLE 1
"Ham" Group and "Turkey-Ham" Group Beliefs

Product Attributes	Ham Beliefs (N=200)	Turkey-Ham Group Mean Beliefs (N=200)	T-Test Significance
1. The food product in this package is <u>relatively nutritious</u> for people to eat.	4.72	4.66	.75
2. The package for this food product meets all federal government <u>safety standards</u> for food packaging.	6.16	6.25	.47
3. The food product in this package is made <u>only from "pork"</u> meat.	4.00	1.74	.001
4. The food product in this package contains <u>more than one type of meat</u> .	4.02	4.75	.001
5. The food product in this package <u>contains some "turkey"</u> meat.	1.63	5.24	.001
6. The food product in this package contains <u>both "turkey" and "ham"</u> meat.	1.70	4.31	.001
7. The food product in this package is made <u>only from "turkey"</u> meat.	1.16	3.35	

Consumer Attribute Beliefs Not Supported by Factual Evidence

Attribute: More Than One Type of Meat. After their exposure to the ham label, respondents in the "ham" group tended to believe (\bar{X} -4.02) that the product they examined did not contain more than one type of meat product. On the other hand, respondents who were exposed to the "turkey-ham" product label tended to believe that this product did contain more than one type of meat product (\bar{X} -4.75).

In general, the factually "correct" answer on the seven-point scale for respondents in both the "ham" and "turkey-ham" groups is "1" since, in fact, both of the representative products contain only one type of meat (i.e., "ham" and "turkey" meat respectively). Stated differently, respondents in each group who perceive correctly that the product package and label they have examined has only one meat ingredient should have indicated a low number on the seven-point scale (i.e., a higher number on the seven-point scale would suggest that a consumer is likely to believe that the product has more than one meat product as an ingredient).

In the present data, the statistically significant lower belief scores for the "ham" group respondents--(compared to "turkey-ham" group respondents)--suggest that ham group respondents tended to correctly perceive that the ham product is made only of one meat product. Conversely, the statistically higher mean belief values for turkey-ham group respondents (compared to ham group respondents) suggest that these consumers believe the turkey-ham product is made of more than one meat product. Since ham group respondent beliefs (evolving from their exposure to the ham product label) agree with "fact,"

while turkey-ham group respondent perceptions do not agree with "fact," it appears that consumers are deceived and misled by the "turkey-ham" wording in the label in both package size categories.

In summary, it appears consumers are deceived in their beliefs about the "more than one type of meat" attribute of the turkey-ham product--based on the exposure to the label on the product. Consumers are not deceived in their beliefs about the "more than one type of meat" attribute of the ham product--based on their exposure to the label on the product.

Attribute: Both "Turkey" and "Ham" Meat. As expected, the "ham" group respondents do not believe that the ham product contains both turkey and ham meat since the mean belief value is extremely low (\bar{X} =1.70). On the other hand, the "turkey-ham" respondent group mean value (\bar{X} =4.31)--a significantly higher belief value than the ham group--suggests that respondents believe the turkey-ham product to contain both turkey and ham meat.

This finding suggests that ham group respondents, based on their exposure to a ham product label, correctly perceive that ham does not contain both turkey and ham meat. Conversely, the significantly higher mean belief score from the turkey-ham group respondents suggest that there respondents believe that turkey-ham contains both ham and turkey meat--a perception which does not correspond to the actual ingredient mix of the product. In other words, the "correct" answer for respondents being exposed to both product labels should be "1" on the seven-point scale. This pattern is much more characteristic of "ham" product group respondents than "turkey-ham" group respondents.

In summary, it appears that consumers are deceived and misled by the "turkey-ham" wording on the label in both package size categories. Consumers are deceived in their beliefs about the "turkey and ham" attribute of the turkey-ham product. Consumers are not deceived in the beliefs about the "turkey and ham" attribute of the ham label product.

Attribute: "Turkey" Meat Only. The "ham" group respondents produced an extremely low mean belief score (\bar{X} =1.16) when asked if ". . . the food product in this package is made only from turkey meat." However, the low mean belief value produced by "turkey-ham" group respondents (\bar{X} -3.35)--which is significantly higher than the ham group belief values--suggests that consumers are deceived by their product label for this attribute.

The "correct" answer for the "ham" respondent group is, in fact, a "1" on the seven-point scale since the ham product does not contain any turkey meat. Conversely, the "correct" answer for the "turkey-ham" respondent groups is, in fact, a "7" on the seven-point scale since the turkey-ham product contains only turkey meat. In essence, the very low group beliefs for the ham group show consumers have a correct perception of reality. On the other hand, the higher belief mean for the turkey-ham group should have been much closer to the upper limit (i.e., a "7" on the scale). In other words, "turkey-ham" group respondents do not believe that turkey-ham is made only from turkey meat, when in fact, this is the only ingredient of the product.

In summary, it appears that consumers are deceived and misled by the turkey-ham wording on the label in both package size categories. Consumers are deceived in their beliefs about the "turkey only" attribute of the turkey-ham product--based on their exposure to the label on the product. Consumers are not deceived in their beliefs about the "turkey only" attribute of the ham product--based on their exposure to the label of the product.

Consumer Attribute Beliefs Supported by Factual Evidence

Attribute: Nutrition. Based on their exposure to the package label, "ham" group respondents ($\bar{X}=4.72$) did not have a statistically significant higher belief than the "turkey-ham" group respondents ($\bar{X}=4.66$) that ". . . the food product in this package is relatively nutritious for people to eat."

Since the mean scores representing consumer beliefs (i.e., evolving from the label) about the nutritional value of the products based on exposure to the labeling are correct in terms of technical evidence, it is reasonable to conclude that consumers in both groups were not misled or deceived by the product labels. In other words, both products are, in fact, relatively nutritious to eat and most consumer correctly perceive this fact.

In summary, it appears that consumers are not deceived in their beliefs about the "nutritional" value attribute of the two products--based on their exposure to ham and turkey-ham labels.

Attribute: Federal Government Safety Standards in Packaging. There is no statistically significant difference between beliefs held by ham group ($\bar{X}=6.16$) and turkey-ham group ($\bar{X}=6.25$) respondents on the attribute that ". . . the package for this food product meets all federal government safety standards for food packaging" based on their exposure to the package label.

Consumers in both groups hold equally strong and equivalent beliefs that the packaging technology--which represents federal government standards--results in "safe" (in terms of purity or contamination) products. Further, all belief data is "high" on the seven-point scale which is "correct" in a normative sense. In summary, consumers are not deceived in their beliefs about the "packaging safety" attribute of two products--based on their exposure to ham and turkey-ham labels.

Attribute: "Pork" Meat Only. The mean value ($\bar{X}=4.00$) of consumers exposed to the "ham" label was significantly higher than the mean value ($\bar{X}=1.74$) of the turkey-ham group respondents.

Since the "ham" products are made only from pork meat and respondent belief scores show recognition of this fact (i.e., a relatively "high" belief score) and since turkey-ham products are not made only from pork meat and respondent beliefs are consistent with this fact (i.e., with relatively low belief scores) it is reasonable to conclude that consumers in both groups were not misled or deceived by the product labels. In essence, it appears consumers are not deceived in their beliefs about the "pork meat only" attribute of the two products--based on their exposure to ham and turkey-ham labels.

Attribute: Some "Turkey" Meat. The belief scores of the "turkey-ham" group respondents ($\bar{X}=5.24$) are significantly higher than the belief scores for the "ham" group respondents ($\bar{X}=1.63$) for the attribute that ". . . the food product in this package contains some turkey meat."

The consumer belief data, in terms of the "turkey-ham" and "ham" group scores on the seven-point scale, reflects an acceptable pattern of results in terms of marketplace reality. In other words, respondents who were exposed to the turkey-ham product label tend to believe there is actually turkey meat in the product (i.e., producing a mean belief value in excess of "5" on the seven-point scale). Similarly, respondents who were exposed to the ham product label tend to believe that there is no turkey meat in the product (i.e., producing mean value near "1" on the seven-point scale). These data suggest that consumers are not deceived in their beliefs about the "turkey" meat attribute of the two products--based on their exposure to ham and turkey-ham labels.

General Food Knowledge Data. Less than one percent (.7 percent) of the sample respondents felt that the term "ham" meant a "cut of any animal," while a majority (51.7 percent) felt the term "ham" referred to a "cut of pork meat" and another (28.5 percent) felt this term referred to a "cut of ham meat." Further, only 3.2 percent felt ham referred to the "smoked and cured thigh of any animal," while 11.2 percent said it was a "smoked and cured thigh of swine." These data suggest that the term "ham" refers to pork or ham meat and very few believe this term applies to a "cut from any animal" or the thigh portion of any animal.

For general comparative purposes, it should be noted that 78.6 percent of the respondents correctly associated the term "hot dog" or "frankfurter" with beef and/or pork products. Very few (.5 percent) of the respondents felt these products are made of poultry products. Similarly, 91.9 percent of the respondents correctly perceived that a "chicken hot dog" is made: (a) only of chicken, or (b) some combination of chicken and beef or pork.

Sociodemographic Data of Respondents. The following statements summarize the results from the sociodemographic questions included in the study. It should be noted that post-experimental analysis of these data indicated that there was no statistically significant difference between "ham" group and "turkey-ham" group respondents in terms of any of those characteristics.

1. About one-third of the sample (31.5 percent) were between 25 to 34 years old, while 17.0 percent and 16.0 percent were between 35 to 44 years old and 45 to 54 years respectively. Persons in the under 25 year-old age group (15.5 percent), 55 to 64 year-old group (12.5 percent), and 65 and older (8.0 percent) accounted for smaller segments in the sample.
2. One-half (50.0 percent) of the respondents came from families with annual incomes between \$10,000 and \$19,999. On the other hand, about one-quarter (23.5 percent) of the families in the sample had incomes in excess of \$20,000 per year, while 20.8 percent of the sample families had incomes less than \$10,000 per year.
3. Nearly all of the sample respondents (90.2 percent) were the head of their household (31.0 percent) or their spouse (59.2 percent). Nearly two-thirds (64.0 percent) of the respondents were female.
4. Almost one-half (49.0 percent) of the respondents were from families with 2 to 4 members, while one-quarter (25.0 percent) of the respondents were from two person families. Few of the respondents were from one person (9.7 percent) or 5 or more person families (16.0 percent).
5. Nearly one-half (49.0 percent) of the respondents were high school graduates and 21.2 percent had some college experience. About one-fifth (18.2 percent) held a college or advanced/professional degree.
6. Over one-half (52.0 percent) of the sample were employed--on a full-time (33.5 percent) basis or as part-time (18.8 percent) employees.
7. About three-quarters (73.0 percent) of the respondents were white and 11.0 percent were non-white. Less than one-tenth (6.3 percent) were Spanish-speaking.

Summary and Implications

Data from this study suggest three summary findings. First, consumers incorrectly believe--after examining an actual product package and label--that "turkey-ham" products: (1) contain more than one type of meat, (2) contain both "turkey" and "ham" meat, and (3) do not contain only "turkey" meat. This causal relationship between exposure to the "turkey-ham" wording on the product label and these incorrect consumer beliefs suggest that this label is a source of "claim-fact" discrepancy deception. On the other hand, consumer beliefs about "turkey-ham" nutritional, packaging, pork meat content, and turkey meat content attributes appear to correctly reflect available scientific evidence. Second, consumers who were exposed to "ham" product label wording did not exhibit false beliefs about any of the seven product attributes measured in the study. Third, study data indicate that very few consumers correctly perceive the technical definition of "ham" used by the USDA. The implications of these findings are presented in terms of the two major study objectives.

Conceptual and Methodological Implications

Three major implications related to research in consumer deception are suggested by this study. First, the study results indicate--on a preliminary basis--that selected conceptual and methodological perspectives from the existing research tradition in advertising deception are applicable to the study of potentially deceptive label wording. In other words, falsely held consumer beliefs associated with specific product attributes represented in product label wording were identified by the study methodology.

Second, future research should extend the findings from this study in two ways. First, additional research on potentially deceptive label wording in other consumer product categories should be conducted to provide documentation of the present study results. Second, other forms of marketing stimuli which may represent sources of marketplace deception--sales catalog product and pricing descriptions, point of sale materials, consumer brochures, etc.--should be investigated. In this way, knowledge on the sources and impact of consumer deception will be broadened in a meaningful way.

Third, the present study illustrates that previous laboratory-based approaches to the study of deception can be applied to "real world" products and consumers in a field experiment context. This factor illustrates that increasing degrees of external validity in research on deception can be achieved by consumer researchers. As such, the credibility of consumer research input public policy decision-making on issues involving deception is strengthened. In stating these consumer research implications the authors recognize also that future research must overcome the major limitations of the present study (i.e., non-probability sampling processes were used at four subjectively selected interviewing locations; testing effects related to exposure to products in a single category; "one time" analysis *vis-a-vis* longitudinal analysis, etc.).

Public Policymaking Implications. Findings from this study suggest that USDA should review its policy position of allowing the "turkey-ham" wording on the label of products in this category since consumer beliefs about the ingredient or content attributes of these products are incorrect. Within the framework of existing policy alternatives at USDA, a post-market regulatory strategy--whereby this wording is removed from the label--would seem to be an appropriate remedy. Conversely, USDA labeling policies for "ham" products do not stimulate incorrect attribute beliefs, and as a result, it appears that no regulatory action affecting the wording on these labels is needed.

These two post-market policy implications are supported by two additional long run considerations. First, since many new combination or multi-ingredient meat products are introduced in the marketplace each year, USDA should consider developing a continuing and on-going program of consumer research in product labeling using: (1) experimental methodologies, and (2) belief variables. Such a research program would gather empirical data on product labeling (in all USDA jurisdictional product categories) in support of pre-market regulatory policies at USDA. The objective of such a program would be to reduce the likelihood of consumer deception emanating from product label wording by preventing such labeling from reaching the marketplace.

Second, since consumers appear to relate the term "ham" to pork and/or ham meat products, the USDA should review the appropriateness of its dictionary-based definition within the context existing consumer knowledge and terminology. The alternative policy approach at USDA--which does not seem to be cost-effective based on findings in this study--would involve a large-scale consumer education program designed to promote a consistent relationship between the public policy definition and consumer understanding of the term "ham." In any event--the preceding discussion of post-market, pre-market, and consumer education implications emanating from this study highlight three potential roles consumer research can fulfill in public policymaking processes in agencies like the USDA.

References

- D. A. Aaker, "Deceptive Advertising," in D. A. Aaker and G. S. Day, eds., Consumerism: Search for the Consumer Interest, 2nd ed. (New York: The Free Press, 1974), 137-156.
- G. M. Armstrong and J. P. McLennan, "The FTC and the Investigation and Regulation of Deception in Advertising," in Thomas V. Greer, ed., Increasing Marketing Productivity/Conceptual and Methodological Foundations of Marketing (Chicago: American Marketing Association, 1973), 430-434.
- G. M. Armstrong and F. A. Russ, "Detecting Deception in Advertising," MSU Business Topics, 23, No. 2 (Spring, 1975), 21-31.
- M. T. Brandt and I. L. Preston, "The Federal Trade Commission's Use of Evidence to Determine Deception," Journal of Marketing, 41(January, 1977), 54-62.
- D. Cohen, "The Federal Trade Commission and the Regulation of Advertising in the Consumer Interest," Journal of Marketing, 33(January, 1969), 40-44.
- _____, "Surrogate Indicators and Deception in Advertising," Journal of Marketing, 36(July, 1972), 10-15.
- T. Dillon, "What is Deceptive Advertising?" Journal of Advertising Research, 13(October, 1973), 9-12.
- R. F. Dyer and P. G. Kuehl, "The 'Corrective Advertising' Remedy of the FTC: An Experimental Approach," Journal of Marketing, 38(January, 1974), 48-54.
- _____, "The Effectiveness of the FTC's 'Corrective Advertising' Policy: New Empirical Insights," in Sharon K. Zeigler, ed., Proceedings of the 1974 National Conference of the American Academy of Advertising (Boston: American Academy of Advertising, 1974), 166-177.

G. T. Ford, P. G. Kuehl and O. Reksten, "Classifying and Measuring Deceptive Advertising: An Experimental Approach," in Proceedings: 1975 Fall Educators' Conference of the American Marketing Association, Rochester, New York. (pages unassigned)

D. M. Gardner, "Deception in Advertising: A Conceptual Approach," Journal of Marketing, 39(January, 1975), 40-46. (Based on "Can Deceptive Advertising be Measured?" Working Series in Marketing Research, No. 18. The Pennsylvania State University, College of Business Administration, June 1973. 22 pages.)

J. E. Haefner, "The Legal Versus the Behavioral Meaning of Deception," in M. Venkatesan, ed., Proceedings of the Association for Consumer Research, Fall, 1972, 356-360.

J. E. Haefner, M. T. Birchmore and S. E. Permut, "Applications of Belief Measures to Specific and Overall Advertising Claims: A Program for Research," in S. K. Zeigler, ed., Proceedings of the 1974 National Conference for University Professors of Advertising, American Academy of Advertising, 311-317.

J. E. Haefner and S. E. Permut, "An Approach to the Evaluation of Deception in Television Advertising," Journal of Advertising, 3, No. 4, 40-45.

H. Keith Hunt, "Source Effects, Message Effects, and General Effects in Counteradvertising," in M. Venkatesan, ed., Proceedings, 3rd Annual Conference, Association for Consumer Research, 1972, 370-381.

_____, "Effects of Corrective Advertising," Journal of Advertising Research, 13(October, 1973), 15-24.

J. Jacoby and C. Small, "The FDA Approach to Defining Misleading Advertising," Journal of Marketing, 39(October, 1975), 65-68.

M. G. Jones, "The FTC's Need for Social Science Research," Proceedings, 2nd Annual Conference, Association for Consumer Research, 1971, 1-9.

H. H. Kassarian, C. J. Carlson and P. E. Rosin, "A Corrective Advertising Study," in Mary Jane Schlinger, ed., Advances in Consumer Research, Vol. 2, 1975, 631-642.

P. G. Kuehl and R. F. Dyer, "Brand Belief Measures in Deceptive-Corrective Advertising: An Experimental Assessment," in Kenneth L. Bernhardt, ed., Proceedings of the 1976 American Marketing Association Fall Conference, 1976, 373-379.

_____, "Applications of the 'Normative Belief' Technique for Measuring the Effectiveness of Deceptive and Corrective Advertisements," in William D. Perreault, Jr., Advances in Consumer Research, Vol. 4, 1977, 204-212.

Michael B. Mazis and Janice E. Adkinson, "An Experimental Evaluation of a Proposed Corrective Advertising Remedy," Journal of Marketing Research, XIII (May, 1976), 178-183.

J. A. Miller, D. A. Topel and R. E. Rust, "USDA Beef Grading: A Failure in Consumer Information?" Journal of Marketing, 40(January, 1976), 25-31.

J. T. Rosch, "Marketing Research and the Legal Requirements of Advertising," Journal of Marketing, 39(July, 1975), 69-79.

"United States of America Before Federal Trade Commission in the Matter of Warner-Lambert Company," Docket No. 8891, Final Order, December 9, 1975 (103 pages).

W. L. Wilkie, "Research on Counter and Corrective Advertising." Paper presented at the American Marketing Association Conference on Advertising and the Public Interest, Washington, D.C., May 9-11, 1973, 16 pages.

W. L. Wilkie and D. M. Gardner, "The Role of Marketing Research in Public Policy Decision-Making," Journal of Marketing, 38(January, 1974), 38-47.

EVALUATION OF FOOD LABELING POLICIES
THROUGH MEASUREMENT OF CONSUMER UTILITY

Roger Best, University of Arizona
Jim McCullough, University of Arizona

ABSTRACT

Food labeling continues to be a controversial topic among consumers, nutritionists, manufacturers, educators, and agencies of the government. The difficulties that have been experienced in past in terms of food labeling awareness, usage, and understanding as well as the consumers right to know more about the composition of food products will effect future food labeling policy decisions. The method of conjoint analysis presented in this paper offers a way of evaluating consumer preferences for existing and alternative labeling policies and method of inferring the degree of satisfaction that would be derived under alternative policies by any given segment of consumers as well as by the market in general.

INTRODUCTION

Underlying public policy decisions concerning food product labeling is the assumption that greater quantities of more detailed information will permit consumers to make improved brand choice decisions (French and Barksdale, 1974). Motivation for better food labeling was outlined at the 1969 White House Conference on Food, Nutrition, and Health, and the United States Food and Drug Administration (FDA) has subsequently required additional label information. Current policy requires one of the following formats for food products.

1. Ingredient and Nutrition Labeling: For products which have been fortified with additional nutrients and/or food products making a nutritional claim.
2. Ingredient Labeling: For all other food products. (Nutrition labeling of these products is voluntary but if disclosed must follow the format prescribed by the FDA).

Under this policy labeling formats have had limited success. Consumer understanding and usage of label information has been disappointing. Nevertheless, there are many who favor even more detailed disclosure of food product composition and nutrition. Existing as well as proposed labeling alternatives are complex requirements which are likely to present difficulties for food manufacturers, technologists, and consumers. The objectives of this study were to: 1) discuss current and proposed food labeling with respect to consumer usage, understanding, and ability to process food labeling information, 2) empirically measure consumer utilities for existing and proposed changes in food labeling, and 3) infer the degree of consumer and overall market satisfaction with alternative food labeling policies.

FOOD LABELING: ALTERNATIVES, USAGE, & DIFFICULTIES

Ingredient labeling found on most food products provides little more than content information. Listings of ingredients in decending order based on quantity is not sufficient for purposes of making an accurate brand choice decision in many instances. For example, individuals on restricted salt diets need to know the amount of salt in a bread product, not simply that there is less salt in the bread than flour, water, and sugar. In response to this type of need, the FDA is considering a

percentage ingredient labeling requirement for most packaged foods and has already proposed percentage ingredient labeling for baby food.

The requirement for percentage ingredient labeling is not a simple change in policy and will require specification of a standard format. If the policy dictates percentage ingredient labeling, the FDA must decide whether the policy should require listing the percentage of only the "characterizing" ingredient or ingredients or the percentage of all ingredients? Or, should the policy require a format listing percentage ingredients down to some cutoff value such as 1% or 2%, after which the other ingredients would be listed in decending order? Whatever the exact format, this requirement must also take into account problems created by this type of labeling in terms of production and quality control as well as from the standpoint of consumer information processing. Percentage ingredient labeling will increase manufacturing costs and these costs will almost certainly be passed on to consumers, adding to already increasing food prices. Imposition of these costs must be reconciled with the increased benefits provided by a percentage ingredient label and the net welfare or satisfaction of all consumers directly affected by such a requirement.

Nutrition labeling has been required for the past three years for those food products to which nutrients have been added or which make a nutritional claim. However, despite the fact that most agree that nutrition labeling was an important step toward presenting consumers with better information for making brand choice decisions (Beloian, 1973; USFDA, 1974b), nutrition labeling has been a controversial policy that has been disappointing in many ways (Daly, 1976).

A number of studies, including some by the FDA have indicated that consumers want nutrition labeling, claim they would use it, but seldom actually use it or understand it correctly in brand choice decisions (Lenahan et. al., 1973; USFDA, 1974a; Daly, 1976; Day, 1976). In an early study Lenahan and his associates found that only 25% of consumers they studied were aware of the nutrition label, 15% understood it, and 10% used it in a purchase decision. An FDA study found that 33% of the consumers they interviewed looked for ingredient labeling and only 5% looked for nutrition labeling. In another earlier study, Asam and Bucklin (1973) reported that the greatest impact of detailed nutrition labeling was in a more favorable perception of product quality. In a more recent review article, Day (1976) has noted that the only apparent effect of nutrition labeling has been to increase consumer confidence in nutritionally labeled food products. The value of nutritional labeling is questioned even further by Peterson's (1976) study of consumer perceptions of nutritional value as a function of nutrition labeling, price, and color of the product in which he found that the color of a bread product was more important than price and nutritional labeling combined in determining consumer perceptions of nutritional value.

Regardless of these findings, the view held by some (Bymers, 1972) is that "the use the consumer makes of nutritional information is peripheral to the main issue of the consumers right to know." Proponents of this view advocate policies leading to increased volume and

detail of food labeling information. This raises a more pragmatic question: Do consumers have the capacity and ability to process this type of information load?

Lambert (1976) evaluated the question of capacity by taking 57 branded food products that could be considered in making up a single days menu. Using linear programming he processed the information such that the selection of brands would meet minimum daily nutritional requirements while minimizing the cost of the purchase. Facing some 54,000 pieces of information it took an IBM Model 370/165 Computer system 7 1/2 minutes of computer time to determine an optimal selection of brands. Since shoppers are faced with even greater purchasing tasks, it seems incomprehensible that the average consumer has capacity to process such large volumes of complex information accurately. With respect to choice accuracy Jacoby and his associates (1974a, 1974b) suggest that providing the consumer with increasing amounts of information may lead to a problem they have termed "information overload." When the consumer becomes overloaded with information his brand selection becomes dysfunctional and the accuracy of his choice is diminished. Considering all of these factors it seems quite possible that policies aimed at increased disclosure of the composition and nutritional value of food products may actually increase the cost of food products and lessen the ability of consumers to make improved brand choice decisions. While the question of choice accuracy under varying types and amounts of food label information has not been resolved, the purpose of this study was to measure and evaluate consumer utility for alternative labeling formats and amounts of information.

METHOD AND ANALYSIS

In this study variations in label format and information load were used to create 15 different bread labels. Bread was selected as a stimulus since it represents a familiar product and one for which the existing labeling format is well standardized. As shown in Table 1 the format of the bread labels was varied in five ways to present labels as ingredients only, ingredient plus nutrition, nutrition only, percentage ingredient plus nutrition, and percentage ingredient labeling. The information load was varied in three ways by constructing

Presenting these labeling stimuli to a consumer and obtaining their rank order preference for these different labels provides the information needed for monotonic analysis of variance, or conjoint analysis of preference data. Given the assumption that consumer utility for alternative formats and amounts of information is an additive function, this analysis of preference data provides part-worth utility functions for each component of an additive utility function. Thus, from a consumer's ranked preference for alternative labels created by variations in format and amount of information, the following utility function can be derived separately for any given consumer:

$$U = U_F + U_L$$

Where:

U = a consumer's overall utility for a given label.

U_F = a consumer's part-worth utility for a given food label format.

U_L = a consumer's part-worth utility for a given amount of labeling information (i.e., information load).

These utilities can be determined individually and used to identify different consumer segments in the market (Assael, 1976) on the basis of similar utility functions. Recognizing the idiosyncratic nature of utility functions representative of different consumer segments can aid a policy makers ability to understand the labeling preferences of any given segment, recognize the diversity of preferences for food labeling represented in the market, and infer the satisfaction or dissatisfaction of any given consumer segment as well as the overall market with alternative food labeling programs.

One-hundred and forty female household shoppers were selected at random from the telephone directory of a city with a population of around 500,000. Although this sampling frame slightly reduces the number of high and low income consumers included in the sample, no attempt was made to compensate for this bias.

Each shopper was interviewed in a 20-minute long in-home interview. During the interview the shopper was presented seven food labels that varied in terms of format and amount of information. The food label shown in

TABLE 1
VARIATIONS IN LABEL FORMAT AND INFORMATION LOAD THAT WERE USED AS STIMULI

LABEL FORMAT	INFORMATION LOAD ⁽¹⁾		
	4 ITEMS	8 ITEMS	12 ITEMS
Ingredients	2-Ingredients	6-Ingredients	10-Ingredients
Ingredient Plus Nutrition	1-Ingredient 1-Nutrition	3-Ingredients 3-Ingredients	5-Ingredients 5-Ingredients
Nutrition Only	2-Nutrition	6-Nutrition	10-Nutrition
Percent Ingredient Plus Nutrition	1-% Ingredient 1-Nutrition	3-% Ingredients 3-Nutrition	5-% Ingredient 5-Nutrition
Percent Ingredient Only	2-% Ingredients	6-% Ingredients	10-% Ingredient

(1) Each label also indicated a price of 54¢ and a net weight of 24 ounces. These were considered as 2 items of information.

labels which contained either 4, 8, or 12 items of information. In each case two of the items of information were price and net weight. Other items were added to the label in the order they would normally appear on an actual label.

Figure 1 is typical of the food labels presented to shoppers. An overlapping Latin-Square design was used to reduce the number of stimuli that needed to be evaluated by the consumer yet retain enough information about alternative formats and amounts of information such that preferences for these stimuli could be

FIGURE 1
A TYPICAL STIMULUS SHOWING A FORMAT OF
NUTRITION AND INGREDIENTS AND AN
INFORMATION LOAD OF EIGHT ITEMS

Price 54¢
NUTRITION
62 calories/slice
12 gm carbohydrates/slice
1.33 mg protein/slice
INGREDIENTS
Flour
Water
Sugar
Net Weight 24 oz.

processed and part-worth utilities derived. Without this procedure it is quite likely that the consumer would be overburdened with this particular type of evaluative task.

The shoppers were asked to evaluate each of the seven bread labels they were presented and then rank the labels in the order they would prefer them as labels they would use in evaluating a brand of bread. After ranking the labels, the participants were instructed to use a Likert-type scale to evaluate each label on the basis of several attitudinal statements which are shown in Table 2. Following this they were asked to evaluate several more general attitudes that were thought to be common

functions for label format and amount of information were aggregated using a cluster analysis program (Diehr, 1974). Discernable clusters were then evaluated in terms of differences in attitude, demographics, or consumptive characteristics that might be unique to consumers with like preferences for bread labeling.

RESULTS AND DISCUSSION

The results of this consumer evaluation task and mathematical derivation of part-worth utilities for alternative food label formats and amounts of information provided three very distinct consumer groups. This is best seen in Figure 2 by graphs of the part-worth utility functions for each of the three consumer groups.

Group I (N = 35) was termed "label avoiders." These 35 consumers provided preferences which were analyzed and converted to part-worth utilities that showed a decreasing utility for more information and a great deal of utility for existing labeling whether it be ingredient or ingredient plus nutrition labeling. The fact that these consumers prefer a bread label with essentially two items of labeling information (shown in Figure 3) led us to infer from their part-worth utility for information load that they may in fact prefer no labeling, just price and net weight. These consumers were heavy users of white sandwich bread and predominantly blue collar families. With respect to other consumers, however, these consumers held attitudes that were more optimistic about how other consumers used food labels as shown in Table 3.

Group II (N = 36) was termed "information seekers."

TABLE 2
ATTITUDES RELATED TO A CONSUMERS PREFERRED FOOD LABEL

ATTITUDE STATEMENT	PERCENT GIVING FAVORABLE RESPONSE IN GROUPS		
	I	II	III
This information is <u>very easy</u> to understand.	76	99	98 ⁽¹⁾
This information would be <u>useless</u> in making a bread purchase decision.	18	21	16
This information <u>fully discloses</u> the quality of this product.	48	45	58
I would <u>not</u> have much faith in this information.	42	33	35
The majority of this information is simply <u>not</u> necessary.	21	12	29
This type of information is <u>needed</u> to determine the real value of a bread product.	76	87	87
The information presented on this label is <u>easy</u> to read.	98	91	99

(1) High and low values are significantly different at $p < .10$.

among bread purchasers (these are shown in Table 3). Finally, the participants were asked several demographic questions relating to the occupation of the major wage earner, family structure, and questions concerning their consumption of bread products.

Ninety-nine usable questionnaires were obtained. There was no evidence that respondents whose results could not be analyzed were different than the 99 consumers examined. Individual preference rankings were analyzed using MONANOVA (Kruskal and Carmone, 1968) and part-worth utility function derived for each of the 99 consumers. Consumers with like or similar part-worth utility

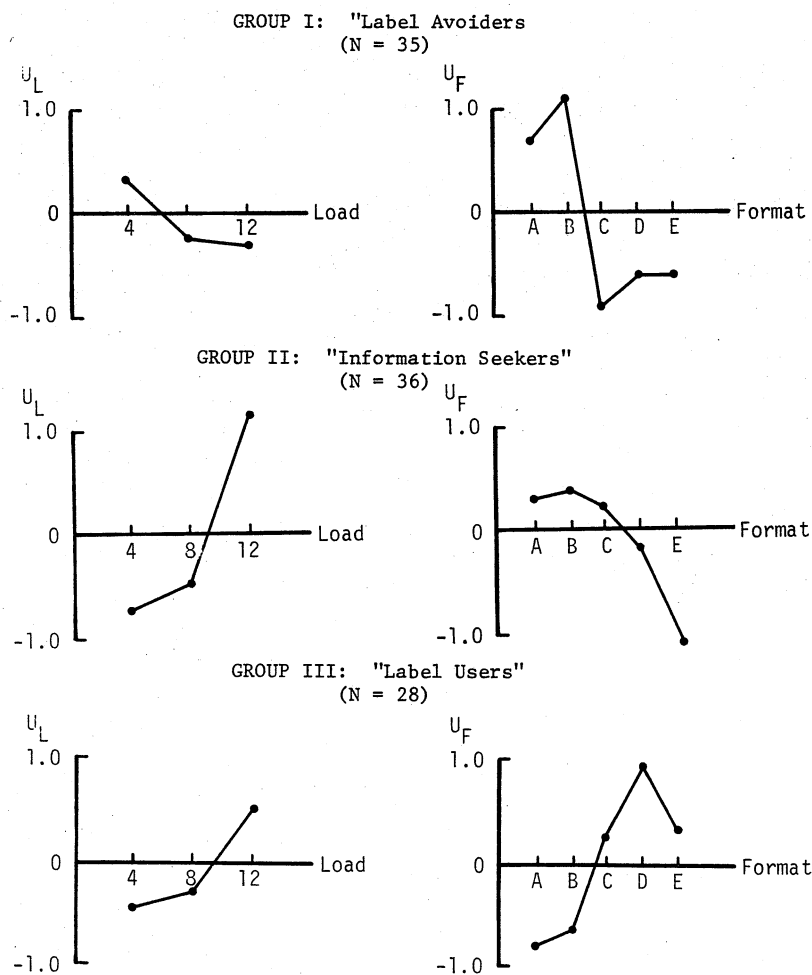
These 36 consumers exhibited a very strong preference for increasing amounts of information. However, they would be most satisfied with the existing format of labeling information as their utility for more detailed information diminished very acutely as shown in Figure 2. Thus, given this overwhelming preference for more information with existing format the food label in Figure 3 for the group termed "information seekers" depicts the label this group preferred the most. This group of consumers were predominantly white collar workers and consumed very little white sandwich bread, but considerably more whole wheat bread, and were more pessimistic about other consumers use of food labels.

TABLE 3
GENERAL ATTITUDES TOWARD BREAD LABELING

ATTITUDE STATEMENTS	PERCENT GIVING FAVORABLE RESPONSE IN GROUPS		
	I	II	III
Shoppers really do <u>not</u> pay much attention to labeling information.	57	72	67 ⁽¹⁾
Nutritional labeling is more <u>useful</u> than ingredient labeling.	60	50	57
Most bread shoppers just look at the <u>price</u> and size of the loaf.	66	72	81 ⁽¹⁾
Most people are concerned with <u>ingredients</u> in bread.	34	39	33
Bread producers try to <u>hide</u> ingredient information on the package.	29	31	30
Most shoppers do <u>not</u> have the knowledge to interpret the information on most labels.	69	56	63
I approve of the use of preservatives in the bread I purchase.	37	33	41

(1) High and low values are significantly different at $p < .10$.

FIGURE 2
PART-WORTH UTILITY FUNCTIONS FOR ALTERNATIVE LABELING FORMAT* AND INFORMATION LOAD



*A = Ingredient
B = Ingredient plus nutrition
C = Nutrition only

D = Percent ingredient plus nutrition
E = Percent ingredient only

FIGURE 3
PREFERRED FOOD LABEL FOR EACH CONSUMER GROUP

GROUP I
N = 35
u = 1.50

Price 54¢
INGREDIENTS Flour Water
Net Weight 24 oz.

GROUP II
N = 36
u = 1.59

Price 54¢
NUTRITION 62 calories/slice 12 gm carbohydrates/slice 1.33 mg protein/slice .44 mg iron/slice 5.3 mg calcium/slice
INGREDIENTS Flour Water Sugar Suet Yeast
Net Weight 24 oz.

GROUP III
N = 28
u = 1.45

Price 54¢
NUTRITION 62 calories/slice 12 gm carbohydrates/slice 1.33 mg protein/slice .44 mg iron/slice 5.3 mg calcium/slice
INGREDIENTS 77% Flour 22% Water .5% Sugar .2% Salt .13% Yeast
Net Weight 24 oz.

Group III (N = 28) was termed "label users." This consumer segment preferred greater amounts of information in greater detail. This would be a group least satisfied with existing labeling requirements and the one most likely to be satisfied with an FDA move to percentage ingredient plus nutrition labeling. Based upon their part-worth utilities for more and detailed labeling information these consumers were termed "label users" and their preferred food label based on what was presented in this study is shown in Figure 3. Members of this group were also predominantly white collar workers, moderate to heavy users of white sandwich bread, and also pessimistic about other shoppers use of food product labeling as shown in Table 3.

The analysis of preference data and its value in identifying discernable differences in consumer utility for alternative food labels could be very useful in evaluating the satisfaction of specific consumer groups as well

as overall market satisfaction under alternative labeling policies. In dealing with the food labeling question, for example, examination of the part-worth utilities for various labeling alternatives among different consumer groups such as those shown in Figure 2 can provide a clear indication of the consumer satisfaction in each of these groups for policy decisions that alter the amount and/or type of labeling information provided to these consumers.

Since it is possible to segment markets on the basis of differing utilities for labeling information, it is then possible to infer the approximate satisfaction or utility each segment of the market would obtain under different food labeling policies. To make this type of inference the following measure of consumer satisfaction can be used to determine labeling satisfaction for any given consumer or consumer segment in the market.

$$CS_i = \frac{U'_F + U'_L}{U_F + U_L}$$

Where:

CS_i = consumer satisfaction of consumers in the i^{th} segment.

U_F, U_L = consumer utility for their preferred food labeling format and information load, respectively.

U'_F, U'_L = consumer utility for existing or proposed food labeling format and information load, respectively.

Market satisfaction can then be inferred by aggregating the satisfaction of each consumer segment with respect to the market share associated with the relative size of each consumer segment as shown below.

$$MS = \sum_{i=1}^N (MS_i) (CS_i)$$

Where:

MS = overall market satisfaction.

MS_i = market share of i^{th} consumer segment; $i=1,2,\dots,N$.

CS_i = consumer satisfaction of consumers in the i^{th} segment.

Utilizing these measures of satisfaction, consumer and market satisfaction with alternative bread labeling programs can be inferred for the consumer groups identified in this study. For any given consumer segment there exists a combination of format and information load that will provide consumers in that segment with maximum satisfaction. When a labeling policy corresponds to a consumer's preferred format and information load, consumer satisfaction will be maximum and indexed as 1.00 using the formula presented above. Labeling programs that are less desirable will result in less consumer satisfaction and a consumer index of satisfaction less than one. In some cases when a labeling program offers a format and information load that are essentially the inverse of the consumer's utility for labeling, the index for consumer satisfaction will take on a negative value. Weighting the consumer satisfaction index of each consumer segment by the relative size of each segment leads to an inference of overall market satisfaction. It also is 1.00 when maximized and can take on negative values under labeling programs that a majority of consumers would find very undesirable.

Utilizing these formulas and the part-worth utility functions shown in Figure 2, the following analysis is provided to illustrate the satisfaction or dissatisfaction of each consumer segment and overall market satisfaction under three different food labeling policies.

Program 1: Ingredient Labeling

A policy of ingredient labeling would correspond to a food labeling program for bread that would provide

ingredient labeling on six ingredients. With the addition of price and net weight this corresponds to an information load of eight items. Consumer satisfaction for each consumer segment and overall market satisfaction under this type of program could be inferred as:

$$CS_I = \frac{.79 - .16}{1.14 + .36} = 0.42$$

$$CS_{II} = \frac{.37 - .48}{.41 + 1.18} = -.07$$

$$CS_{III} = \frac{-.76 - .13}{.91 + .54} = -0.61$$

$$MS = (.354)(.042) + (.364)(-.07) + (.282)(-.61) = -.05$$

Program 2: Ingredient plus Nutrition Labeling

In this study this type of bread labeling policy would correspond to a format of ingredient and nutrition labeling with an information load of 12 items. Consumer and overall market satisfaction under this policy could be inferred as:

$$CS_I = \frac{1.14 - .20}{1.14 + .36} = 0.63$$

$$CS_{II} = \frac{.41 + 1.18}{.41 + 1.18} = 1.00$$

$$CS_{III} = \frac{-.64 + .54}{.91 + .54} = -0.07$$

$$MS = (.354)(.63) + (.364)(1.00) + (.282)(-.07) = .56$$

Program 3: Percentage Ingredient plus Nutrition Labeling

This type of bread labeling policy would require (on the basis of this study) 12 items of information that were in a format of half percentage ingredients and half nutritional information. Under this program consumer and overall market satisfaction could be inferred as:

$$CS_I = \frac{-.53 - .20}{1.14 + .36} = -0.49$$

$$CS_{II} = \frac{.07 + 1.18}{.41 + 1.15} = 0.78$$

$$CS_{III} = \frac{.91 + .54}{.91 + .54} = 1.00$$

$$MS = (.354)(-.49) + (.364)(.78) + (.282)(1.00) = .39$$

Based on these inferences a food labeling policy of ingredient plus nutrition labeling would provide the greatest market satisfaction of the three programs evaluated. For consumers in groups I and II their satisfaction was greatest with this type of labeling policy; group II in fact would have maximized their satisfaction for bread labeling under this type of program.

Ingredient labeling produced the least overall market satisfaction, yet it is a permissible food labeling alternative for food products not containing nutrients or making a nutritional claim. Except for consumers in group I all other consumers would be considerably dissatisfied with this type of bread labeling policy.

Program 3, a policy of percentage ingredient plus nutrition labeling would achieve a moderate overall market satisfaction while maximizing the utility for consumers in group III. However, overall market satisfaction with this type of bread labeling policy is hindered by consumers in group I who prefer less information in a simpler format. If the consumers in group I actually prefer no labeling other than price and weight, or are simply indifferent toward labeling format, then the consumer satisfaction inferred from group I is misleading and should be discounted on the basis of no interest in bread labels. If this were the case, then program 3

would provide the greatest satisfaction for bread labeling among consumers represented in this study. On the other hand, if consumers in group I prefer less information in a simpler format because they do not use it but realize that a greater quantity of more detailed information will lead to higher bread prices, then their utilities for simpler bread labels, and subsequent satisfaction under alternative programs can not be ignored in evaluating overall market satisfaction under these programs.

CONCLUSIONS

This study examined consumer responses to variations in bread labeling format and information load using consumer preferences for alternative bread labels and analysis of these preferences using the method of conjoint analysis. Application of partial preference analysis to identify differences in consumer responses offers a useful method for assessing the impact of policy decisions on the consumer and market in general. Analyzing the cost of policy decisions can be made with relative ease in monetary terms, but assessment of its impact on the consumer satisfaction can be considerably more difficult. The methodology presented in this paper offers some benefit to those needing to analyze the latter effect of alternative policy decisions.

This type of analysis is particularly useful as a segmentation tool that differentiates the market on the basis of utility. With respect to food labeling this procedure allows researchers and policy makers to evaluate existing and alternative labeling programs with regard to the idiosyncratic preferences of different consumer segments, recognize the diversity of these preferences, and infer the likely satisfaction of dissatisfaction of any given consumer segment and overall market under any specific food labeling policy. In some instances it may be possible to specify different labeling requirements based upon differences in product usage in different consumer segments of the market. Though this type of marketing program would contribute to greater consumer and market satisfaction, it is probably not feasible in most situations.

In this study we have demonstrated how preference data could be analyzed using the method of conjoint analysis and used as a tool for evaluating alternative food labeling programs. Though the results in many ways concur with previous studies, additional research is planned to validate the results reported here. In the next two stages of our research we plan to sequentially: 1) evaluate in a choice situation the choice consistency between a consumers labeling preference and the brand of bread selected when only the labeling format is varied, and 2) evaluate the degree to which a consumer with differing bread labeling preferences will pay for additional costs of information. Beyond this additional research needs to focus on the relationship between consumption and preference as well as the understanding and use of labeling on more complex products such as over-the-counter drugs, ointments, and remedies.

Food labeling will continue to be a controversial topic among consumers, nutritionists, manufacturers, educators, and agencies of the government. The difficulties that have been experienced in past in terms of food labeling awareness, usage, and understanding as well as the consumers right to know more about food products will effect future food labeling policy decisions. The method of conjoint analysis presented in this paper offers a way of evaluating consumer preferences for labeling under existing and alternative policies and a method for inferring the degree of satisfaction that would be derived under policy alternatives by any given segment of consumers as well as by the overall market in general.

REFERENCES

- E.H. Asam and L.P. Bucklin, "Nutrition Labeling for Canned Goods: A Study of Consumer Response," Journal of Marketing, 37 (April 1973), 32-37.
- H. Assael, "Segmenting Markets by Response Elasticity," Journal of Advertising Research (1976).
- A. Berloian, "Nutrition Labels: A Great Leap Forward," FDA Consumer (September 1973), 1.
- G. Byrners, "Seller-Buyer Communication: Point of View of a Family Economist," Journal of Home Economics, 64 (February 1972), 59.
- P. Daly, "The Response of Consumers to Nutrition Labeling," Journal of Consumer Affairs, 10 (Winter 1976), 170-78.
- G.S. Day, "Assessing the Effects of Information Disclosure Requirements," Journal of Marketing, 40 (April 1976), 42-52.
- G. Diehr, "Cluster Analysis Program (Revised," College of Business Administration, University of Washington, (1974), mimeographed.
- W.A. French and H.C. Barksdale, "Food Labeling Regulation: Efforts Toward Full Disclosure," Journal of Marketing, 38 (July 1974), 14-19.
- J. Jacoby, D. Speller, and C. Kohn, "Brand Choice as a Function of Information Load: Replication and Extension," Journal of Consumer Research (December 1974), 68-72.
- J.B. Kruskal and F.J. Carmome, Jr., "MONANOVA, a FORTRAN IV Program for Monotone Analysis of Variance," Bell Telephone Laboratories (1968), mimeograph.
- A.L. Lambert, "Nutrition Information: A Look at Some Processing and Decision-Making Difficulties," in W.D. Peneault, Jr., ed., Advances in Consumer Research, 4 (Fall 1976), 126-32.
- R.J. Lenahan, J.A. Thomas, O.A. Taylor, D.L. Call, and D.I. Padberg, "Consumer Reaction to Nutritional Labels on Food Products," Journal of Consumer Affairs, 7 (Summer 1973), 1-14.
- C.S. Martinsen, J.G. Ostrander, and J.M. McCullough, "Consumer Attitudes Toward Preservatives in Bread and Other Foods," Bakers' Digest (1975).
- R.A. Peterson, "Consumer Perceptions as a Function of Product Color, Price and Nutrition Labeling," in W.D. Peneault, Jr., ed., Advances in Consumer Research, 4 (Fall 1976), 61-63.
- U.S. Food and Drug Administration, "Consumers Talk about Labeling," FDA Consumer (February 1974a)
- U.S. Food and Drug Administration, "The Food Labeling Revolution," FDA Consumer (April 1974b)

CONSUMER PROCESSING AND USE OF SUPPLEMENTARY DRUG LABEL INFORMATION

F. Robert Dwyer, University of Arizona

Abstract

Right to know arguments, comprehension effects on compliance and shortcomings of verbal communications from medical staff have fostered increased interest in written patient-oriented information on drugs. This paper presents an experimental approach for the evaluation of cognitive and behavioral impacts of alternative communication designs.

Introduction

The extent of patient non-compliance with prescription drug therapy has been measured at rates ranging from 4 to 90%, with the average near 35 to 45% (Boyd *et al*, 1974). The consequences of patient overdoses, omissions, and improper mixing include deprived therapeutic benefits and distorted evaluation of therapy in addition to economic wastage.

Studies of the causes of patient non-compliance have identified several patient and treatment factors--perceived seriousness of the illness, side effects from the medication, confidence in the physician. Blackwell (1973) has reviewed the literature on compliance and concludes, however, that beyond patient and treatment factors

The most important contribution to compliance is the understanding a patient has of the illness, the need for treatment and the likely consequences of both. (p. 252)

Moreover, this point of view has been supported by recent research (Clinite and Kabat, 1976; Cole and Emmanuel, 1971; and Madden, 1973). The issue is not so clear-cut, though, as Malahy (1966) and Sackett *et al* (1975) failed to find compliance effects from increased comprehension of treatment and illness.

Nevertheless, educating patients about their prescription drugs is becoming an important component in health care. Fueled by this trend has been a growing interest in the subject of patient package inserts (PPI's)--written patient-oriented information on drugs. Recently a measure (HR 14289) was introduced before the House that would require PPI's be distributed with all prescription drugs. Another, S-1282, requiring PPI's for drugs undergoing the final phases of clinical evaluation was proposed by Kennedy in the Senate. Not unexpectedly, a good deal of political controversy accompanying the PPI interest hinges on the patient's right to know about his or her condition and treatment, versus the physician's responsibility to create a favorable clinical milieu, which is sometimes assumed to require the withholding of information on medication side effects and interactions.

Beyond the "right-to-know" stance, additional arguments advanced by PPI proponents include the shortcomings of the physician's verbal counseling of the patient: (1) verbal counseling is often forgotten, (2) the physician may be too busy to counsel, and (3) the patient's perception of his or her doctor's time-pressure may discourage patient questioning. Verbal counseling by the pharmacist has similar problems. In addition, the prescription may be dispensed to a third party (a person sent to the pharmacy by the patient), and the pharmacist who

prepares the prescription may not be the person who physically distributes the medication.

PPI opponents counter with arguments that inserts will lead to dangerous attempts of patients to self medicate. Also, the provision of cautions and enumeration of even rare side effects may foster undue anxieties in patients. Obviously, cost and distribution problems are also raised.

Fortunately, at a recent symposium in Washington, D.C., PPI proponents and opponents recognized the inadequacies of current research and agreed on the importance of addressing the following key questions:

1. What ought patients to know about the medicines prescribed for them?
2. What effect will such information have on patients and on the practice of medicine?
3. How can the information be best communicated?
4. What is the best delivery method? (Schmidt, 1976)

The purpose of this article is to address the above questions by highlighting significant findings and shortcomings of previous research in the area. Moreover, the literature review leads to a set of hypotheses that were tested in a factorial field experiment run on consumers of anti-hypertensive diuretics.

Previous Research and Experience

Several studies have examined the question of what information should be provided to patients. Smith (1975) and Hussar (1975) recommend patients be informed of the (1) purpose of the medication, (2) proper times and methods of administration, and (3) special precautions to be observed while taking the medication. Joubert and Lasagna (1975) assessed the information needs of consumers via telephone survey and found the name of the drug, its purpose, and adverse effects and cautions in normal use desired by 97, 93, and 89%, respectively. Similar percentages were obtained in a Food and Drug Administration survey of oral contraceptive users (FDA, 1975).

A recent study by Katz and Rose (1976) suggests that a reliance on survey data for assessing consumer information needs can be misleading. A majority of their respondents wanted additional information on light bulbs that, in fact, had been heretofore provided for several years. In proper perspective, it seems that asking consumers if they would like to know more about the drugs they take may be equivalent to asking them if steps should be taken to lower property taxes.

The research on the effect of amount of information on decision making may have PPI significance. As a source of health education information, PPI's should help consumers make several kinds of decisions--the decision to initiate, alter, or discontinue drug therapy or contact with the medical staff, and the "implicit" decision a patient makes every time he or she remembers or forgets to take a prescribed drug (Morris, 1977).

In simulated war games and lesser developed country economy management games, Suedfeld and Streufert (1966)

and Schroder, Driver, and Streufert (1967) consistently discovered an inverted U-shaped relationship between optimal decision making and amount of input information. Recently Jacoby and associates (1974; 1974b) have found evidence of possible overload effects in a simulated purchase situation. Although several scholars have challenged their interpretations (Summers, 1974; Wilkie, 1974a), and the external validity of the brand X attribute matrix is suspect, the overload phenomenon is nevertheless intriguing.

Indeed, the PPI concept is just one significant dimension of the accelerating interest in consumer information disclosure in a number of product categories (Wilkie, 1974b; Day, 1976). Therefore, a good deal may be learned by examining previous experiences in disclosure measures.

Field experience in non-drug information disclosure has shown little evidence of overload ("confused, dysfunctional") behavior, rather, minimal to nonuse of the policy-provided information. The interpretation of these results has primarily occupied two dimensions: (1) the compatibility of the newly disclosed information with the abilities of consumers to process (attend, comprehend, believe, and retain) this information, and to a lesser extent, consumer motivation to use the information. A brief examination of the performance of recent disclosure laws will serve to highlight these points.

The results of unit pricing experiences have shown consumer awareness to vary between 50 and 82% (Monroe and LaPlaca, 1972; Carman, 1972-73). While 5 to 38% claimed they had used the information to switch package sizes or change brands, warehouse movements reflected no such shifts. Houston (1972) bypassed the possibility that price is often a nondeterminant factor by instructing his subjects to shop for the most economical brand and size. Here, when consumers were motivated to shop "economically" unit price savings were achieved.

Before leaving the topic of unit pricing, it is worthwhile to note that a portion of the variance in consumer-reported familiarity and use of unit price information can be explained by the differences in the manner in which the information was presented. Formats have included shelf tags, computer print-outs, and display signs. Following several laboratory studies of decision making, Russo *et al* (1975) hypothesized that a single list of all brand/sizes and their unit prices is an effective presentation of unit price information. In a five-week sales test on three products there was a significant shift in sales from higher to lower priced brands within each brand subgroup (local/national).

In another area, Day (1976) argues that truth-in-lending (TIL) requirements suffer from design problems. While finding just 10% of credit buyers of durables claiming use of the TIL information, he suggests the buyer does not get the correct information when it is needed, prior to the point of decision, but after the effective oral agreement has been concluded.

Finally, an experiment by Asam and Bucklin (1973) is of considerable significance because it not only shows that attention varies by format but that consumers make product inferences that vary with the amount of detail or preciseness of (nutritional) information.

While these brief reviews of non-drug disclosure laws have served well to illustrate the importance of presenting the "right" information, in the "right" format, at the "right" level of detail, at the "right" time, perhaps even greater insights can be gleaned from the lone PPI evaluation.

Since 1970 the FDA has required that basic warnings and instructions be distributed with birth control pills. In addition, two brochures, one prepared by the American

Medical Association and the other by drug manufacturers, providing a fuller discussion of risks and benefits were distributed by the physician upon the patient's request. In 1975 the FDA contracted for a nationwide survey of former and current oral contraceptive (OC) users to collect information about distribution, readership, comprehension and reported behavioral changes induced by the inserts.

The results of this study (FDA, 1975) have been recently summarized by Morris *et al* (1977). The survey disclosed that 93% reported receipt of the PPI and 94% of those read it. Only 54% knew that OC's could cause blood clots, in spite of the fact that the inserts stated that blood clots were the most serious side effect of the pill. Also, only one-third of the OC users who reported having never received an insert were "concerned" about not getting one. The non-readers most frequently said they "always ask (their) doctor such questions."

The AMA and manufacturers' brochures were reported less widely distributed. Yet, among those who received the brochure, an equally high (95%) level of readership was reported. Interestingly, more women retained the often more detailed manufacturer-prepared brochure (50%) than the AMA brochure (16%). Only 11% report consulting either PPI or brochure on occasions subsequent to the initial prescription receipt. The brochure and insert were consulted upon instances of side effects, missed pills, or when the physician was unavailable.

Twelve percent of the current users who had read the insert reported that it had raised questions that caused them to contact their physicians. About one in seven insert readers report the insert causing either increased or decreased visits or calls to the doctor. Two-thirds of these decreased physician contacts. Among the brochure readers about the same proportion report physician contact changes. In four-fifths of the cases here, however, the action was toward reduced physician contact.

At least two policy implications from this research have been identified by the FDA. First, since patients tended to physically retain the more detailed brochures, it has been proposed that the brochure be expanded to include more detailed information related to pill use. Secondly, since the insert had nearly three times the distribution level of the physician-distributed brochure, the FDA has proposed the pharmacist directly distribute the pills and insert.

While the oral contraceptive insert/brochure experience has provided a great number of insights for the development of a full-scale PPI program, . . .

OC users are not typical of the general patient population and survey research cannot answer all critical questions. Studies of patients undergoing treatment for disease are also needed to help determine the best methods of communicating important patient information (Morris *et al*, 1977, p. 25).

Hence, Morris *et al* complete the rationale for the study described below.

Methodology

An input-output approach was employed on newly diagnosed hypertensive patients in a 3² factorial design, in a clinical setting, to determine what presentation format and what level of detail in the information contained in the insert were optimal in terms of:

1. Enhancing compliance and clinical outcome
2. Facilitating patient "processing" of the information

3. Fostering optimal levels of anxiety

Specifically, the design was employed to test the following set of hypotheses:

1. Form of presentation and level of detail of PPI information have no effect on patient compliance and clinical outcome.
2. Form of presentation and level of detail of PPI information have no effect on patient perceived experiences with side effects.
3. Form of presentation and level of detail of PPI information have no effect on patient treatment-related anxiety.
4. Form of presentation and level of detail of PPI information have no effect on patient information processing (attention, yielding, comprehension) of drug information.
5. Form of presentation and level of detail of PPI information have no effect on patient physical retention of the insert.
6. The coefficients of the likely covariates of compliance, experience with side effects, information processing, and physical retention of the PPI are all zero. These covariates include:
 - a. Attitude toward physician
 - b. Concern for health
 - c. Perceived seriousness of condition
 - d. Sex
 - e. Education, occupation, etc.
 - f. Health Locus of Control score

Treatments

Russo and company's (1975) unit pricing experiment proves the importance of the form of presentation in consumer information disclosure programs. Also, by following Bettman's (1975) line of thought in information environment design, it should be kept in mind that consumption of medication over time means possible adverse reactions or complications over time, requiring recollection of information from the medical staff or PPI. Unless the patient tape records conversations with medical personnel, the feature of external information storage is unique to the PPI. It seems desirable, then, to attempt to develop a PPI that would facilitate a patient's physical retention of it.

It is possible, however, that the creation of a widely retained insert is an insurmountable (patients misplace or dispose of a designed-to-be-retained insert) or unnecessary task (patients hold on to almost anything). If the former is the case, the focus of PPI design efforts should be on gaining and holding the patient's initial attention and interest in its contents.

In the event that neither a designed-to-be-retained insert nor the designed-to-be-attractive insert offer significant improvements in patient comprehension, retention and compliance, or if patients are motivated to attend to nearly any format, the design focus should be on minimizing the cost of the PPI.

It follows, then, that the form of presentation of drug information, henceforth designated factor A, should be tested at a minimum of three levels. The specific factor A levels are summarized in Table 1.

The information categories appropriate for patient package inserts, the literature suggests, are essentially settled. Yet to be determined, however, is the relative amount of detail an insert should provide in each of the categories.

TABLE 1
Factor A: Form of Presentation

Level	Description	Rationale Summary
A ₁	Thin, onion skin paper	Minimum cost
A ₂	6-panel, 2-fold colorful brochure	Attract immediate attention
A ₃	Page for prototype "family medical guide"	Easy reference (designed to be retained)

"Amount of detail" is a versatile phrase that can encompass a number of dimensions. Amount of detail on side effects, for example, could include gradations within the areas of (1) descriptive information, (2) remedies, or (3) causes. In this study, gradations in amount of detail have been made across the areas enumerated above. Specifically, the levels progress from "what can happen," to "what to do if it does," to "why it happens." This sequence is best defended by paraphrasing McGuire (1976): People already know side effects, risks, and medication errors are bad; people must be taught how they can cope with and avoid them.

The "whys" of side effects and precautions may have one or more of the following effects: (1) enhance/hinder the believability of the PPI, (2) soothe anxieties accompanying the medication program, or (3) foster anxiety through confusion. Therefore, their inclusion or exclusion from the PPI is a significant policy issue.

The amount of detail, henceforth designated factor B, was examined at the levels summarized in Table 2.

TABLE 2
Factor B: Amount of Detail

Level	Description	Rationale Summary
B ₁	What can happen, do's, don'ts	Descriptive information only (Minimum detail)
B ₂	B ₁ plus solutions to problems	How to cope (Medium detail)
B ₃	B ₂ plus why problem arises	Effects on believability, comprehension (High detail)

The design of the test inserts was facilitated by the information presented in the Physician's Desk Reference and Govoni and Hays (1971). Insert prototypes were reviewed by the head pharmacist and three doctors on staff at the supporting medical center. Inserts were modified in light of their remarks.

Type size was held constant across all treatments, while a graphics consultant was employed in the design of the front cover of the brochure insert, A₂. Sixty-pound green construction paper was the vehicle for A₂ and A₃ was bound in a green Duo Tang notebook. The amount of detail, factor B, levels are depicted in the Appendix.

Experimental Design

The experimental design employed in this study provided for replicates of a 3². Experimental subjects in this design were newly diagnosed hypertensive patients at the St. Louis Park Medical Center in the Minneapolis suburb. Furthermore, to facilitate cost control and minimize variance in the treatment regimen, the population of

eligible subjects was limited to essential hypertensives prescribed diuretics. Finally, due to limitations of experimenter control, only patients receiving their prescriptions at the Medical Center Pharmacy were eligible for study. The director of the Medical Center had estimated that 30% of the center-prescribed medications were dispensed at the Medcenter Pharmacy. There was no evidence indicating that those not dispensing there were different in regards to the outcome measures to be discussed below.

Inserts in coded bags were dispensed at the pharmacy by a double blind procedure. Two weeks after the receipt of the medication and insert, the patient received a letter from the pharmacist. The letter stated that unless the patient called to state otherwise a researcher from the University of Minnesota would soon call to set up a half-hour interview dealing with his or her opinions on medications, health, and health care in general. About one in six patients did not wish to be called.

The patient interview was completed approximately one month after medication and insert receipt. Patient comprehension of the insert information was assessed in Section 1, a set of 20 true-false, yes-no, and multiple choice questions. Two of the 20 questions were dummy items on topics not discussed in any insert but included to confound yea-saying.

Section 2 of the questionnaire explored the patient's recollection of receipt of the insert (aided and unaided). In addition, physical retention and recalled use of the insert were examined. Insert believability or patient yielding was measured with an itemized rating scale nested among measures of PPI complexity and completeness. Also, patients' valuations of other sources of drug information were obtained.

Patient recollection of side effects was solicited with an open ended question. The interviewer probed for the frequency of occurrence of each experienced side effect and whether or not the patient took special actions--such as calling the medical staff, ceasing medication, etc.

Section 3 employed a lead-in and probing to assess patient compliance with medication directions. This section also obtained patients' prior familiarity with the drug and family experience with hypertension.

Section 4 measured several suggested covariates of compliance and was completed by respondents by hand. Importance and belief scores were obtained on five attributes of physicians in general and the prescribing physician. These attributes, identified in patient focus groups, were: (1) knows his/her business, (2) is friendly, (3) communicates in patient terms, (4) charges reasonable rates, and (5) shows real concern for the patient.

A second portion of Section 4 consisted of items from the A-State anxiety measure developed by Spielberger *et al* (1970). This scale has been used to measure the level of anxiety raised by a situation, as opposed to the A-Trait measure which seeks to assess the anxious tendencies of individuals. The instrument has never been used to measure the anxiety caused by a communication message and should be interpreted with caution; it is possible that the interview itself fosters anxiety in individuals. Also, since it was felt that the original 20-item instrument was too threatening and fatiguing for the study purpose, a random selection of 10 items yielded a more workable instrument. Nested among the 10 items were three similar statements designed to assess subjects' degree of concern for their health.

The third portion of Section 4 was a health locus of control measure developed by Wallston *et al* (1975).

This scale has been used to examine an individual's belief in internal control of physical health or well being. In their validation efforts, the developers of this scale noted some evidence of greater information seeking among subjects classified as internals.

Finally, respondents indicated their occupation, level of education, marital status, and number of children (under 18 years) living at home.

In addition to the data obtained from patient interviews, subject deviations from the prescription refill date and changes in blood pressure readings were obtained from Medical Center and Pharmacy archive data. Blood pressure readings used were those taken closest to one month after the initial diuretic prescription was dispensed.

Results

At this point in time, data has been obtained on 30 subjects--providing approximately three replicates. This apparently small sample size is characteristic of previous research in patient education and is indicative of the difficulties encountered in such field research. Nevertheless, the data contain a number of interesting facets and are presented as preliminary in light of continuing data collection.

First, to briefly contrast this experiment with the previously reported OC experience, 100% reported receipt of the PPI on diuretics. Physical retention of the insert by hypertensives was higher (67% vs. 50%) but the OC retention rate was measured much later from the time of initial receipt. Forty-eight percent of the hypertensives (vs. 11% of the OC users) reported looking back at the insert. Most common reasons for subsequent insert consultations were to recall special cautions and side effects.

Subjects scored an average of 69% on the measure of comprehension; the most commonly missed items dealt with the drug's effects on pregnant and breast-feeding women. Overall, the inserts were rated clear and helpful. In addition, subjects rated the inserts on a par with the prescribing physician in terms of contribution to their understanding the nature of the drug.

A two-factor fixed effects analysis of variance/covariance model was examined for each of the dependent variables enumerated in hypotheses 1-6. Since there were unequal cell sizes the dummy variable regression approach was employed as covariate coefficients, interactions and main effects were tested at $\alpha = .1$. The results of these tests are addressed in the following paragraphs and are summarized in Table 3.

Two dependent variables were employed in testing hypothesis 1, which dealt with treatment effects on compliance. The number of days delinquent (early) from the compliant refill date is the variable labeled "Refill deviance" in the first row of Table 3. While there are no main effects on this measure of compliance, patients who received minimum information on onion skin were significantly more deviant.

Hypothesis 1 was also tested on a dichotomous criterion, patient self report of compliance. Here the hypothesis is rejected on the basis of significant factor B effects. In examining the coefficients from Table 3 two items deserve special notice. First, among factor levels, the largest increment in compliance comes in B_2 , in which remedies and strategies for coping are provided. Interestingly, this information was also included in B_3 , which did not foster increased compliance. Second, and related to hypothesis 6, education beyond high school is positively associated with compliance. No treatment effects on clinical outcome were uncovered.

TABLE 3
Summary of Significant Standardized Regression Coefficients for the Tested Hypotheses

Hypothesis Number	Dependent Variables	Independent Variables								Statistics						
		Factor A		Factor B		Interactions				HCON-CERN*	SEX*	EDUC*	R ²	F	Sig.	
		A1	A2	B1	B2	AB11	AB12	AB21	AB22							
H ₁	Refill Deviance	--	--	--	--	.459	--	--	--	--	--	--	.21	6.142	.02	
H ₁	Compliance Self-report	--	--	--	.358	--	--	--	--	--	--	--	.480	.39	7.070	.004
H ₂	Side Effects	.503	-.163	-.239	.225	--	--	--	--	--	--	--	.53	5.63	.03	
H ₃	Anxiety	--	--	--	--	--	--	--	--	--	--	--	--	--	--	n.s.
H ₄	Receipt	--	--	--	--	--	--	--	--	--	--	--	--	--	--	n.s.
H ₄	Nonbelief (unaided)	--	--	--	--	--	--	--	.547	--	--	--	.30	9.804	.005	
H ₄	Nonbelief (aided)	--	--	--	--	--	--	--	.655	.249	.335	--	.56	8.856	.001	
H ₄	Comprehension	--	--	--	--	--	--	--	.676	.275	--	--	.51	11.301	.001	
H ₅	Physical Retention	--	--	--	--	--	--	--	--	.488	--	--	.18	5.14	.03	

*Indicates covariates identified in hypothesis 6 were also tested.

The presence of both A and B factor effects on whether subjects experienced side effects requires the rejection of hypothesis 2. Patients receiving the onion skin and/or strategies for dealing with difficulties are more likely to notice side effects. This pattern does not seem to be related to any anxiety-inducing treatment effects as hypothesis 3 is accepted.

Hypotheses 4 and 5, dealing with treatment effects on patient information processing, are accepted in their null form. There were several occasions when, as the only model elements, main effects on insert believability, comprehension, and physical retention were significant. However, this significance disappeared as other, more powerful variables were introduced.

As attention shifts to hypothesis 6 dealing with the impact of the covariates of these processing criteria, a number of interesting relationships are revealed. First, since all experimental subjects recall getting an insert, no significant covariates were expected. However, both aided and unaided measures of insert believability (coded in terms of unbelievability) are inversely correlated with patients' attitudes toward their prescribing physician. In the aided measure, inserts prove less credible to females and patients highly concerned about their health.

Comprehension of the purpose and nature of the drug is positively correlated with patients' attitude toward their doctor and their concern for health. This latter variable is also significant in explaining 18% of the variation in physical retention rates.

Conclusions

The focus of this research has been on the measurement of end-states as they have been effected by the insert stimuli. The information processes themselves were never explicitly measured. However, the alternative approaches, direct monitoring and decision net, become unworkable in field evaluations.

Furthermore, the study does not pretend to have explored all possible vehicles of written drug information. However, the chosen factor levels were of such a diverse nature that if form of presentation effects do exist, they would have been identified in this design. Level of detail effects would seem to be easily manufactur-

able--technical names for side effects and precautions would be nearly impossible for patients to understand. The examined range of this factor was deliberately restricted to what are currently considered realistic levels from a judgmental perspective.

Also, the generalizability of these findings is limited to disease categories similar in nature and consequences to hypertension. It is likely that a patient with an acute condition requires a different insert than the chronic subjects of this study. Further research is required with patients with a variety of acute conditions.

In spite of these limitations, however, several important conclusions seem inescapable. First, on a descriptive level, several insights have been gained into patient use of supplementary drug information. These findings complement the earlier study of OC users. Moreover, the demonstrated effects of form of presentation and level of detail on patient compliance, anxiety, and understanding provide evidence for PPI program and vehicle design.

Regardless of which compliance measure is used, the data suggest that patients need information beyond the basic purpose, side effects, do's, and don'ts. Looking at patient self-reports of compliance, it is clear that strategies for coping with the possible treatment remedies are required. This information may make patients more likely to notice side effects, as does the onion skin format, but no measurable increase in anxiety is fostered by these treatments. It is also worthwhile to note that the provision of B₃ information, the why's, does not enhance compliance, affect experiences with side effects, nor foster insert credibility.

Patient information processing of the inserts is explained primarily by motivational factors--not the communication stimuli themselves. It should be noted, however, that after the motivational covariates, in several instances, factor main effects were only "marginally" insignificant. They may enter the model as the power of the tests increase with continuing data collection.

B₁ Level of Information

YOU AND YOUR MEDICATIONS

YOUR DOCTOR HAS PRESCRIBED A DIURETIC (DI-YU-RET-IK) DRUG. DIURETIC DRUGS ARE MADE BY SEVERAL DRUG COMPANIES. EACH COMPANY HAS ITS OWN NAME FOR THE DRUG IT MANUFACTURES. IT IS IMPORTANT TO KNOW THE NAME OF THE DRUG YOU ARE TAKING. READ CAREFULLY THE INSTRUCTIONS ON THE LABEL. FOLLOW THOSE INSTRUCTIONS.

IT IS IMPORTANT THAT YOU TELL YOUR DOCTOR IF YOU HAVE DIABETES, LIVER OR KIDNEY PROBLEMS, ALLERGY OR BRONCHIAL ASTHMA. ALSO, TELL YOUR DOCTOR IF YOU ARE PREGNANT OR BREAST FEEDING.

THE PURPOSE OF THIS DRUG IS TO REMOVE YOUR BODY'S EXCESS SALTS AND WATER. SOME OF THE DIRECT EFFECTS OF THIS MEDICATION ARE MORE FREQUENT URINATION, OCCASIONAL DIZZINESS, LIGHT-HEADEDNESS OR HEADACHE, AND SOMETIMES DRY MOUTH, THIRST, OR CRAMPS IN MUSCLES.

RARELY, SIDE EFFECTS SUCH AS PERSISTENT FATIGUE OR NAUSEA, STOMACH CRAMPS OR RASH OCCUR. THESE SHOULD BE REPORTED TO YOUR DOCTOR.

FINALLY, YOUR DOCTOR MAY WANT TO SEE YOU AFTER A MONTH OR MORE TO CHECK ON TREATMENT EFFECTIVENESS. CONTACT YOUR DOCTOR SOONER IF YOU HAVE ANY OF THE PROBLEMS MENTIONED ABOVE.

ANY QUESTIONS CAN BE DIRECTED TO THE PHARMACY STAFF BY CALLING:

926-5522

B₂ Level of Information

YOU AND YOUR MEDICATIONS

YOUR DOCTOR HAS PRESCRIBED A DIURETIC (DI-YU-RET-IK) DRUG. DIURETIC DRUGS ARE MADE BY SEVERAL DRUG COMPANIES. EACH COMPANY HAS ITS OWN NAME FOR THE DRUG IT MANUFACTURES. IT IS IMPORTANT TO KNOW THE NAME OF THE DRUG YOU ARE TAKING. READ CAREFULLY THE INSTRUCTIONS ON THE LABEL. FOLLOW THOSE INSTRUCTIONS.

IT IS IMPORTANT THAT YOU TELL YOUR DOCTOR IF YOU HAVE DIABETES, LIVER OR KIDNEY PROBLEMS, ALLERGY OR BRONCHIAL ASTHMA. ALSO, TELL YOUR DOCTOR IF YOU ARE PREGNANT OR BREAST FEEDING.

THE PURPOSE OF THIS DRUG IS TO REMOVE YOUR BODY'S EXCESS SALTS AND WATER. A DIRECT EFFECT OF THIS DRUG IS MORE FREQUENT URINATION. UNLESS OTHERWISE DIRECTED, TO AVOID GETTING UP IN THE NIGHT, TAKE MEDICATION BEFORE NOON.

YOU MAY OCCASIONALLY EXPERIENCE DIZZINESS, LIGHT-HEADEDNESS OR HEADACHE. WHEN THESE OCCUR, REST A WHILE. DO NOT GET UP AND DOWN SUDDENLY. MOVE MORE SLOWLY. YOU MAY GET VERY THIRSTY AND HAVE A VERY DRY MOUTH. FRUITS AND FRUIT JUICES AND FRESH VEGETABLES ARE THE BEST REMEDY FOR THIS DISCOMFORT.

RARELY, SUCH SIDE EFFECTS AS PERSISTENT FATIGUE, NAUSEA, STOMACH CRAMPS OR RASH OCCUR. THESE SHOULD BE REPORTED TO YOUR DOCTOR. REMEMBER, THOUGH, THESE PROBLEMS OCCUR NATURALLY FROM TIME TO TIME EVEN WHEN YOU'RE NOT TAKING ANY MEDICATION. UNLESS THEY OCCUR FREQUENTLY, FOLLOW YOUR USUAL PROCEDURES OF TREATMENT. TAKING THE DRUG WITH MEALS MAY PREVENT STOMACH IRRITATIONS.

FINALLY, YOUR DOCTOR MAY WANT TO SEE YOU AFTER A MONTH OR MORE TO CHECK ON TREATMENT EFFECTIVENESS. CONTACT YOUR DOCTOR SOONER IF YOU HAVE ANY OF THE PROBLEMS MENTIONED ABOVE.

ANY QUESTIONS CAN BE DIRECTED TO THE PHARMACY STAFF BY CALLING:

926-5522

YOU AND YOUR MEDICATIONS

YOUR DOCTOR HAS PRESCRIBED A DIURETIC (DI-YU-RET-IK) DRUG. DIURETIC DRUGS ARE MADE BY SEVERAL DRUG COMPANIES. EACH COMPANY HAS ITS OWN NAME FOR THE DRUG IT MANUFACTURES. IT IS IMPORTANT TO KNOW THE NAME OF THE DRUG YOU ARE TAKING. READ CAREFULLY THE INSTRUCTIONS ON THE LABEL. FOLLOW THOSE INSTRUCTIONS.

IT IS IMPORTANT THAT YOU TELL YOUR DOCTOR IF YOU HAVE DIABETES, LIVER OR KIDNEY PROBLEMS, ALLERGY OR BRONCHIAL ASTHMA -- AS THOSE CONDITIONS CAN BE AGGRAVATED BY THIS DRUG. ALSO, TELL YOUR DOCTOR IF YOU ARE PREGNANT OR BREAST FEEDING BECAUSE THE DRUG CAN BE PASSED TO THE FETUS OR CHILD.

THE AIM OF THE TREATMENT IS TO REMOVE YOUR BODY'S EXCESS SALTS AND WATER. A DIRECT EFFECT OF THE DRUG, THEN, IS MORE FREQUENT URINATION. UNLESS OTHERWISE DIRECTED, TO AVOID GETTING UP IN THE NIGHT, TAKE YOUR MEDICATION BEFORE NOON. DUE TO THE ELIMINATION OF SOME WATER AND SALTS IN YOUR BLOOD, YOU MAY OCCASIONALLY EXPERIENCE DIZZINESS, LIGHT-HEADEDNESS OR HEADACHE. WHEN THESE OCCUR, REST FOR A WHILE. DO NOT GET UP AND DOWN SUDDENLY. MOVE MORE SLOWLY.

YOU MAY GET VERY THIRSTY AND HAVE A DRY MOUTH. THIS IS BECAUSE YOUR BODY NEEDS SOME SPECIAL SALTS. FRUITS AND FRUIT JUICES AND FRESH VEGETABLES ARE THE BEST REMEDY FOR THIS DISCOMFORT.

RARELY, SUCH SIDE EFFECTS AS PERSISTENT FATIGUE, NAUSEA, OR STOMACH CRAMPS OCCUR. THESE MAY BE CAUSED BY AN OVERDOSAGE AND SHOULD BE REPORTED TO YOUR DOCTOR. REMEMBER, THOUGH, THESE PROBLEMS OCCUR NATURALLY FROM TIME TO TIME EVEN WHEN YOU'RE NOT TAKING ANY MEDICATION. UNLESS THEY OCCUR FREQUENTLY, FOLLOW YOUR USUAL PROCEDURES OF TREATMENT. TAKING THE DRUG WITH MEALS MAY PREVENT STOMACH IRRITATIONS. IF A RASH APPEARS, IT MAY MEAN YOU ARE ALLERGIC TO THE MEDICATION. TELL YOUR DOCTOR RIGHT AWAY.

FINALLY, YOUR DOCTOR MAY WANT TO SEE YOU AFTER A MONTH OR MORE TO CHECK ON TREATMENT EFFECTIVENESS. CONTACT YOUR DOCTOR SOONER IF YOU HAVE ANY OF THE PROBLEMS MENTIONED ABOVE.

ANY QUESTIONS CAN BE DIRECTED TO THE PHARMACY STAFF BY CALLING:

926-5522

References

Edward H. Asam and Louis P. Bucklin, "Nutritional Labeling for Canned Goods: A Study of Consumer Response," Journal of Marketing (April 1973), 32-37.

James R. Bettman, "Issues in Designing Consumer Information Environments," Journal of Consumer Research, 2 (December 1975), 169-177.

Barry Blackwell, "Patient Compliance," The New England Journal of Medicine (August 2, 1973), 249-252.

James R. Boyd, Tim R. Covington, Walter F. Stanaszek, and R. Timothy Coussons, "Drug Defaulting. Part I: Determinants of Compliance," American Journal of Pharmacy, 31 (April 1974), 486-491.

James M. Carman, "A Summary of Empirical Research on Unit Pricing," Journal of Retailing (Winter 1972-3), 63-72.

James C. Clinite and Hugh F. Kabat, "Improving Patient Compliance," Journal of American Pharmaceutical Association (January 1976), 74 +.

Phillip Cole and Sister Emmanuel, "Drug Consultation: Its Significance to the Discharged Hospital Patient and its Relevance as a Role for the Pharmacist," American Journal of Hospital Pharmacy (December 1971), 954-960.

George S. Day, "Assessing the Effects of Information Disclosure Requirements," Journal of Marketing (April 1976), 42-52.

- Food and Drug Administration, Survey of Consumer's Perceptions of Patient Package Inserts for Oral Contraceptives (Springfield, Virginia: National Technical Information Service, 1975).
- Laura E. Govoni and Janice E. Hayes, Drugs and Nursing Implications (New York: Appleton-Century-Crofts, 1971), 66-68.
- Lawrence C. Hoff, "How Often Do Consumers Seek Your Advice on Rx and OTC Products," Pharmacy Times (August 1975), 52-55.
- Michael J. Houston, "The Effect of Unit Pricing on Choices of Brand Size in Economic Shopping," Journal of Marketing (July 1972), 51-54.
- Daniel A. Hussar, "Patient Noncompliance," Journal of the American Pharmaceutical Association, No. 4 (April 1975), 183-190 +.
- Jacob Jacoby, Donald Speller, and Carol Kohn, "Brand Choice Behavior as a Function of Information Load," Journal of Marketing Research (February 1974), 63-69.
- Jacob Jacoby, Donald Speller, and Carol Kohn Berning, "Brand Choice Behavior as a Function of Information Load: Replication and Extension," Journal of Consumer Research (June 1974), 33-42.
- Pieter Joubert and Louis Lasagna, "Patient Package Inserts. I. Nature, Notions and Needs," Clinical Pharmacology and Therapeutics, No. 5 (November 1975), 507-513.
- Benjamin J. Katz and Jane A. Rose, "Information Utilization and the Awareness Criterion in Labeling Regulation," 1976 Educators' Proceedings, Kenneth Bernhardt, ed. (Chicago: American Marketing Association, 1976), 202-204.
- Edward E. Madden, Jr., "Evaluation of Outpatient Pharmacy Patient Counseling," Journal of the American Pharmaceutical Association, No. 8 (August 1973), 437-443.
- Bernadine Malahy, "The Effect of Instruction and Labeling on the Number of Medication Errors Made by Patients at Home," American Journal of Hospital Pharmacy (June 1966), 283-292.
- William J. McGuire, "Some Internal Psychological Factors Influencing Consumer Choice," Journal of Consumer Research, 2 (March 1976), 302-319.
- Kent B. Monroe and Peter J. LaPlaca, "What are the Benefits of Unit Pricing," Journal of Marketing (July 1972), 16-22.
- Louis A. Morris, "Patient Package Inserts: A New Challenge or a New Tool?" mimeo (Rockville, Md.: Food and Drug Administration, 1977).
- Louis A. Morris, Michael Mazis and Evelyn Gordon, "A Survey of the Effectiveness of Oral Contraceptive Patient Information," mimeo (Rockville, Md.: Food and Drug Administration, 1977).
- Edward J. Russo, Gene Krieser and Sally Miyashita, "An Effective Display of Unit Price Information," Journal of Marketing (April 1975), 11-19.
- D.L. Sackett, R.B. Haynes, E.S. Gibson, B.C. Hackett, D.W. Taylor, D.W. Roberts, and A.L. Johnson, "Randomized Clinical Trial of Strategies for Improving Medication Compliance in Primary Hypertension," Lancet, 1 (1975), 1205-1207.
- Alexander M. Schmidt, "The FDA Viewpoint," Drug Information Journal, 11 (January 1977), 4-7.
- Harold M. Schroder, Michael J. Driver and Siegfried Streufert, Human Information Processing (New York: Holt, Rinehart and Winston, Inc., 1967).
- Dorothy L. Smith, "Why Should Patients Comply with Medication Instructions?" Drug Intelligence and Clinical Pharmacy (May 1975), 268-269.
- Charles D. Spielberger, R.L. Gorsuch and R. Lushene, Manual for the State-Trait Anxiety Inventory (Palo Alto, California: Consulting Psychologists' Press, 1970).
- Peter Suedfeld and Siegfried Streufert, "Information Search as a Function of Conceptual and Environmental Complexity," Psychonomic Science, No. 10 (1966), 351-352.
- John O. Summers, "Less Information Is Better?" Journal of Marketing Research (November 1974), 461-468.
- Barbara S. Wallston, Kenneth A. Wallston, Gordon D. Kaplan and Shirley A. Maides, "Development and Validation of the Health Locus of Control (HLC) Scale," mimeo (Nashville, Tennessee: Vanderbilt University, School of Nursing, 1975).
- William L. Wilkie, "Analyses of Effects of Information Load," Journal of Marketing Research (November 1974), 262-266.
- William L. Wilkie, Public Policy and Product Information: Summary Findings from Consumer Research, report prepared for the Marketing Science Institute under NSF Grant GI 42037 (Cambridge, Mass.: Marketing Science Institute, 1974).

A MULTIPLE DISCRIMINANT ANALYSIS APPROACH TO THE
DEVELOPMENT OF RETAIL STORE POSITIONING

Lawrence J. Ring, University of Virginia*
Charles W. King, Purdue University

Abstract

This paper applies multiple discriminant analysis to a large-scale consumer data base, to provide an empirical analysis of the factors underlying store choice behavior for men's fashions. These findings define retail store positions in a particular marketplace and pinpoint the major competitive dimensions there.

Introduction

Understanding the clothing and apparel shopping behavior of the contemporary consumer is vitally important to effective marketing in the fashion retail marketplace. Each retailing marketplace is characterized by a wide range of stores and associated strategies and operating methods. The sum total of those strategic and tactical actions taken by a given store or chain of stores is the "product" or "image" which that store presents to the consuming public.

In general, the retailer's objective has been to create a store image which will appeal to a particular segment of the population. It has been widely held [1], [2], [3], and [4] that store "image" is produced by the joint effect of a multiplicity of store or product attributes. The major objective is to determine whether consumers perceive major differences in store appeal when faced with a choice among several (10) competing chains in a given market place. The aim here is to provide an empirical analysis of the factors underlying store choice behavior with respect to men's fashions and to contrast the strengths and weaknesses of several different retailing institutions in the city of Toronto, Canada.

Field Methodology

The data used in this study were gathered as part of a broader on-going research program in Toronto, Ontario, Canada. That program is the Toronto Retail Fashion Market Segmentation Research Program and has been described elsewhere [5] and [6]. The sample used in the analysis here was part of a specially recruited panel of husband and wife pairs in Toronto. A total of 1025 usable questionnaires were received from cooperating male panel members. The questionnaires were mailed to each panel member complete with standard self-administration instructions. Each panel member received a small gift for completing the questionnaire.

In attempting to profile the Toronto men's wear retail marketplace, it was first necessary to define a relevant set of retailers to present to respondents for their judgments across a variety of attitudinal measures.

The 1975 edition of the Toronto Yellow Pages lists over 200 separate retailers under the title "Men's Clothing and Furnishings - Retail" alone -- not including department and discount stores. The vast majority of these

*This study was made possible by funds provided by several Toronto men's wear retailers who wish to remain anonymous, and by research grants to the author from the Colgate Darden Graduate School of Business Administration and the Krannert Graduate School of Management.

stores, however, are relatively small and are single-outlet proprietorships. Therefore, it was expected that a majority of the population would be unfamiliar with stores falling into the single-outlet proprietorship category, and thus unable to comment knowledgeably about their operations.

There are, on the other hand, a small set of stores which most consumers could reasonably be expected to recognize. Toronto's three major department stores and many of its multiple-outlet men's wear chain stores would appear to have sufficient market coverage for respondents to judge them.

In addition, the mass merchandiser (department stores and specialty chain) retailers as a group appear to hold a major (and growing) share of Canadian retail men's wear business, according to the most recent government figures. In 1974, department stores and chain stores obtained 58 per cent of the sales of men's wear in Canada and had maintained an average growth rate for the past five years of 12 per cent dollar sales per year. In contrast, independent stores' dollar sales in men's wear grew only 6 per cent over the same period.¹

For the purposes of this project, ten large chains were mentioned by name (an "any other store" response category was also included to allow respondents to choose stores other than those which were specifically mentioned in the questionnaire).

The detailed store-by-store analysis dealt with fashion retailing as it relates to the major department and major specialty chain stores which are currently actively servicing the Toronto men's wear market. Each of the chains included in the analysis was operating at least six retail outlets in Toronto in 1975. In addition, it was estimated that each of the chains was obtaining gross men's wear sales in Toronto of at least \$5 million dollars per year.

What was not examined in great detail were all the remaining smaller chains and the hundreds of small independent fashion specialty stores. Table 1 represents a comparison of consumers' perceptions of the sample set of chains described above with their perceptions of the aggregated remainder of the stores servicing the Toronto men's wear market.

This comparison reveals that the 10 specific sample chains studied here were well known and held, in the aggregate, dominant positions on all of the dimensions of retail patronage which were measured in the research. On the critical dimensions of "last store shopped" and "store shopped most often", the 10 sample chains were named by approximately 70 per cent of the respondents.

Analysis

The initial analysis pass involved the computation of frequencies of mention across the sixteen measures of patronage which were listed in Table 1. The simple frequency analysis was useful for building a baseline

¹Source: Statistics Canada Documents 63-002 and 63-005, 1974.

statement of individual chain strengths and weaknesses.

Following the computation of frequencies, cross-classification analysis was performed on the determinants of patronage and other key variable sets using the share of shoppers, a surrogate for market share, as the dependent variable. The share of shoppers has historically been measured by market researchers in a variety of ways. In this project, two measures were employed:

1. Which store was visited the last time the respondent bought an article of men's wear?
2. Which store was shopped most often for men's wear?

The purpose of the cross-classification analysis was to pinpoint the strengths and weaknesses of each of the chains as perceived by their own customers for each of the patronage dimensions.

The cross-classification analysis was also useful in the development of decision rules for the grouping of the chains together for further analysis.

The final step in the analysis of the determinants of patronage in the Toronto retail men's wear market involved the construction of the "product space," that is, a spatial representation of the relative positioning of the ten chains included in this survey in Toronto.

This spatial representation was constructed through the use of multiple discriminant analysis. The mathematics of the multiple discriminant method are developed briefly later in the paper.

The objective here was somewhat strategic in nature. The technique permits the identification of potential opportunities in the existing retailing milieu. And, pragmatically, this technique can be useful in discovering directions in which any or all of the existing store images may be modified for greater sales and profitability.

Univariate Results

Univariate cross-classification analysis was used to determine the relative importance of the patronage dimensions to each of the chains covered in this project. For each chain, the cross-classification tables were restructured so that the diagonal values from the cross-classification of each patronage determinant by "store last shopped" were ordered from highest to lowest per cent of mentions. A summary of the diagonals is presented in Table 2. For example, for Chain A², ordering the data from Table 2 highest to lowest, among those who "last shopped" for men's wear at this chain ...

- 69 per cent said Chain A was "easiest to get to from home"
- 36 per cent said Chain A had the "best value for the money"
- 33 per cent said Chain A was the "best for conservative, everyday men's wear"
- 32 per cent said Chain A had the "most knowledgeable, helpful salesclerks"
- 27 per cent said Chain A had the "largest overall merchandise assortment/selection"

- 24 per cent said Chain A had the "most exciting merchandise display"
- 22 per cent said Chain A had the "lowest prices"
- 19 per cent said Chain A had the "best advertising"
- 16 per cent said Chain A was "best for current, up-to-date men's wear"
- 15 per cent said Chain A was "best for very latest, most fashionable men's wear"
- 12 per cent said Chain A had the "highest quality"

Assuming that those dimensions receiving the highest share of mentions for each chain (in this case, Chain A) were, in fact, the most important determinants of retail patronage for that chain, this ordered set of diagonals represented the "retail image" of each chain as seen by the chain's customers.

TABLE 1
Share of Mentions Given to the 10 Chains Surveyed Vs. All Others on Store Shopping Attitude/Determinants of Patronage Questions

	Survey Chains	Any Other Store	Base
Easiest to get to from home	88.9 ^a	11.1%	1021
Last store shopped for men's wear	68.4	31.6	1011
Store shopped most often	71.8	28.2	1014
Lowest prices	91.0	9.0	1010
Highest prices	92.2	7.8	1001
Highest quality	86.6	13.4	1000
Lowest quality	91.4	8.6	993
Best value for the money	88.7	11.3	999
Worst value for the money	83.5	16.5	987
Most knowledgeable, helpful salesclerks	82.8	17.2	994
Largest overall assortment/selection	95.2	4.8	997
Best advertising	91.8	8.2	970
Most exciting merchandise display	83.6	16.4	976
Best for conservative, everyday men's wear	91.5	8.5	998
Best for current up-to-date men's wear	87.9	12.1	972
Best for very latest, most fashionable men's wear	85.0	15.0	988

²Chain names have been disguised for this analysis to protect the proprietary interest of those firms who sponsored the research. Chains A, B, and C are department stores; Chains D and E are discounter/mass merchandisers; and Chains F, G, H, I, and J are men's wear specialty chains.

^aREAD: Among the respondents answering the location question, 88.9 per cent reported one of the 10 sample chains was "easiest to get to." Alternatively, 11.1 per cent reported some other store not among the 10 sample chains was "easiest to get to."

TABLE 2
Summary of Cross Classification Analyses Diagonals for
12 Determinants of Retail Patronage

Determinants of Retail Patronage	Store Last Shopped for Men's Wear									
	Chain A	Chain B	Chain C	Chain D	Chain E	Chain F	Chain G	Chain H	Chain I	Chain J
1. Easiest to get to from home	69%*	42%	34%	33%	42%	9%	50%	24%	4%	32%
2. Lowest prices	22	24	19	92	49	24	18	0	0	16
3. Highest quality	12	18	20	0	0	26	32	83	58	21
4. Best value for the money	36	55	58	39	46	66	52	40	35	47
5. Most knowledgeable, helpful salesclerks	32	35	46	25	27	49	52	89	72	54
6. Largest overall assortment/selection	27	50	51	8	26	39	24	35	15	18
7. Most exciting merchandise display	24	38	32	25	18	36	47	69	59	26
8. Best advertising	19	54	50	0	35	16	20	30	21	10
9. Store shopped most often	58	69	82	60	57	57	57	79	71	45
10. Best for conservative, everyday men's wear	33	54	58	33	27	24	27	28	27	22
11. Best for current, up-to-date men's wear	16	29	29	0	8	28	56	65	85	38
12. Best for the very latest, most fashionable men's wear	15	22	19	0	5	19	36	70	60	24

*READ: Among those who last shopped at Chain A for men's wear, 69 per cent came from Chain A's trading area, 22 per cent said Chain A had the lowest prices, and 12 per cent said Chain A had the highest quality, etc.

Looking at Chain A, for example, the highest percentage was 69 per cent, and it relates to location. This suggests that Chain A's location(s) was a major reason for shopping at this chain. Of lesser importance were "value", "conservative men's wear", and so forth. However, only on location did this chain appear particularly strong.

The univariate analysis of patronage determinants suggested that there were four kinds of chains serving the men's apparel market in Toronto. They were:

1. The discounter/mass merchandisers which appeared to be drawing their customers primarily by their low prices, but also on good value, location, and conservative everyday men's wear.
2. The department stores, which were obtaining their shoppers through good value and conservative, everyday men's wear also, but in addition because of their advertising and large overall merchandise displays.
3. The mid-range fashion specialty chains, which were drawing primarily on value, current, up-to-date men's wear, knowledgeable, helpful salesclerks, and exciting merchandise displays.
4. The high fashion specialty chains, which performed most strongly among their own shoppers on salesclerk help, high quality, very latest, most fashionable and current, up-to-date men's wear, and exciting display.

The question that remains revolves around what is the joint interaction of all the patronage dimensions taken as a whole? Therefore, the final step in the analysis

of the structure of the Toronto retail men's wear market involved the construction of the "product space"; that is, a spatial representation of the relative positioning of the ten chains included in this survey in Toronto.

Multiple Discriminant Mapping

The objective in this exercise was somewhat strategic in nature. Perceptual mapping can be used as a tool for identifying potential opportunities for both the new men's wear retailer and the currently operating retailers. In particular, this technique can be useful to the existing retailer in discovering directions in which to modify his image to produce the greatest sales gain.

In recent years the construction of geometric spacial models of particular markets has received a great deal of attention in market segmentation analysis. And, a number of techniques and combinations of techniques have been employed in constructing these so-called "product" or "brand spaces". Among the most popular tools have been discriminant analysis [9] and [10], and factor analysis [11].

The data gathered for this analysis in the project was obtained through use of the associative techniques. Simply speaking, each respondent was asked to associate one chain with each of the following questions:

1. Which store is the easiest to get to from home?
2. Which store has the lowest prices?
3. Which store has the highest prices?

4. Which store has the most knowledgeable, helpful salesclerks?
5. Which store has the highest quality men's fashions?
6. Which store has the lowest quality men's fashions?
7. Which store gives you the best value for the money in men's fashions?
8. Which store gives you the worst value for the money in men's fashions?
9. At which store do you find the most exciting merchandise display of men's fashions?
10. Which store has the best fashion advertising?
11. Which store has the largest overall merchandise assortment/selection in men's fashions?
12. Which store is best for conservative, everyday men's wear?
13. Which store is best for current, up-to-date men's wear?
14. Which store is the best for the very latest, most fashionable men's wear?

So, essentially each respondent picked "1 of n" (n=11) stores on each of the above dimensions. However, in the course of choosing one chain for each dimension and therefore ranking that chain as the best (or the worst, the largest, and so forth), each consumer assigned a rating to all the stores. That is, in the dummy variable sense, a "1" was assigned to the chosen store and a "0" was assigned to each of the other stores.

As a result, each respondent has rated each chain a "1" or a "0" on each dimension. This, in effect, constitutes approximately 1000 observations across the eleven store alternatives. Given this data, multiple discriminant analysis was chosen as the technique best suited for constructing a spacial model of the Toronto men's wear market.

In the Toronto data, respondents rated a set of eleven men's wear chains on fourteen determinants of patronage. The fourteen dimensions have been considered as the reasons taken together which account for the ways in which Toronto shoppers think men's wear chains differ from one another. The objective is to determine how these "reasons" discriminate among the set of chains.

Results

The results of the discriminant analysis were very consistent with the univariate analysis results presented earlier in this paper. Table 3 reports the structure for the eight most significant discriminant functions and the canonical correlations between the 14 determinants of patronage and the 11 group (chain) membership variables.

The structure coefficients in this table suggest that the first eight discriminant functions appear to be measuring:

1. Low quality/low price (+) versus high price (-)
2. Conservative men's wear with large assortment and advertising (+) versus very fashionable men's wear with high quality and high price (-)

3. High price, large assortment, best advertising (-)
4. Location and high price (-)
5. Worst value and location (+)
6. Very latest, most fashionable merchandise (+)
7. Low price (+) versus low quality (-)
8. Best value for the money (+)

Table 4 indicates that the first two functions accounted for approximately 83 per cent of the discrimination among the 11 Toronto chains. Additional interpretation can be obtained by locating the 11 group (chain) centroids on the discriminant functions.

Figure 1 displays both the group centroids and the determinants of patronage on the first two discriminant functions, and Table 6 presents the locations of the group centroids on the first eight discriminant functions. Here, and in Table 3, the discriminant functions have been standardized in order to adjust for the variability in the original variables and to therefore accurately report the relative importance of the variables.

Figure 1 shows the perceptual map of the Toronto men's wear market in two dimensional form. The location of each store is indicated on these two dimensions. The horizontal dimension seems to contrast low quality and low price on the right with high price on the left. The vertical dimension appears to suggest conservative men's wear with a large assortment and advertising at the top in contrast to very fashionable, high quality, high priced men's wear at the bottom.

The mean rating of each chain on each determinant can be determined by erecting perpendiculars from each centroid to each "determinant vector." For example, Chain H is viewed as the chain with the highest prices, followed by Chain I, "others", Chain G, Chain A, Chain J, Chain F, Chain B, Chain C, Chain E, and Chain D.

Figure 2 displays the territorial map of the Toronto retail men's wear market. In this figure, all space on the graph has been classified according to the store centroid to which it is closest. The territorial map clearly suggest that there are three major chain-type divisions in the Toronto market, the department stores, the specialty stores, and the discounters.

The discriminant analysis indicated that in Toronto the major competitive dimensions of the discounter/mass merchandisers were low quality, low price, and worst value for the money. The fashion specialty chains did best on high price, most fashionable men's wear, high quality, current men's wear, exciting merchandise display, and knowledgeable, helpful salesclerks. The department stores showed strongest on best assortment/selection, best advertising, best conservative everyday men's wear, best location, and best value for the money.

TABLE 3
Determinants of Patronage Correlations and Canonical Correlations for Eleven Chains in the Toronto Men's Wear Store Space

Determinants of Patronage	Canonical Correlation	Discriminant Functions							
		DFI	DFII	DFIII	DFIV	DFV	DFVI	DFVII	DFVIII
		.679	.513	.339	.212	.179	.164	.090	.071
Best Location		-.00297	.12194	.04833	-.42774	.32549	.01339	-.07829	.12558
Lowest Prices		.32591	.00012	-.00951	.01468	-.16307	-.08186	.42513	.06114
Highest Prices		-.14442	-.34183	-.35104	-.28784	-.13587	-.04592	-.03492	.14096
Most Helpful Salesclerks		-.01173	-.02872	.03364	.19265	.24232	-.14698	-.02524	.12044
Highest Quality		-.02429	-.15554	-.18764	.05617	.03572	-.28450	.01871	-.01871
Lowest Quality		.79808	-.04487	-.21587	-.03940	-.02458	.09740	-.28762	.09120
Best Value for the Money		-.03548	.10039	.03273	.19942	-.02370	-.13025	-.01871	.32693
Worst Value for the Money		.07218	-.00027	-.00480	.12660	.33316	-.04067	.03589	-.19967
Largest Assortment/Selection		-.06176	.24891	-.24939	.05969	-.17726	-.01298	-.20943	-.18377
Best Conservative Men's Wear		-.06667	.19337	-.20619	.07192	-.00035	.05670	-.05028	.01073
Best Current Men's Wear		-.01695	-.05750	.06713	.12275	-.06246	.14556	-.18072	.08915
Best Very Latest Men's Wear		-.03369	-.16927	.03188	.12409	.05054	.34305	.09765	.14840
Most Exciting Display		.00448	-.04396	.03011	.01907	.24524	.04522	.16309	-.19879
Best Men's Wear Advertising		-.08026	.24456	-.26539	-.10618	-.01939	.13253	.25559	-.02224

READ: The standardized coefficient of "Best Location" on Discriminant Function I was -.00297. It was .12194 on Discriminant Function II.

TABLE 4
Toronto Men's Wear Market Discriminant Analysis

Discriminant Function	Eigenvalue	Relative Percentage	Canonical Correlation
1	0.85659	58.36	0.679
2	0.35667	24.30	0.513
3	0.12988	8.85	0.339
4	0.04694	3.20	0.212
5	0.03320	2.26	0.179
6	0.02779	1.89	0.164
7	0.00815	0.56	0.090
8	0.00510	0.35	0.071
9	0.00308	0.21	0.055
10	0.00048	0.03	0.022

READ: The eigenvalue of discriminant function number 1 was 0.85659. The first discriminant function accounted for 58.36 per cent of the discrimination among the Toronto chains.

TABLE 5
Toronto Men's Wear Market Discriminant Analysis

Functions Derived	Wilks' Lambda	Chi-Square	Degrees of Freedom	Significance
0	0.3108	13159.898	140	greater than .001
1	0.5770	6191.961	117	greater than .001
2	0.7829	2756.850	96	greater than .001
3	0.8845	1381.675	77	greater than .001
4	0.9261	865.115	60	greater than .001
5	0.9568	497.344	45	greater than .001
6	0.9834	188.698	32	greater than .001
7	0.9914	97.309	21	greater than .001
8	0.9964	40.051	12	greater than .001
9	0.9995	5.458	5	not significant

READ: Before any functions were removed, Wilks' Lambda was 0.3108, indicating that considerable discriminating power existed in the variables being used. The chi-square remains significant until the eighth function has been removed. Only at that point did very little discriminating information remain.

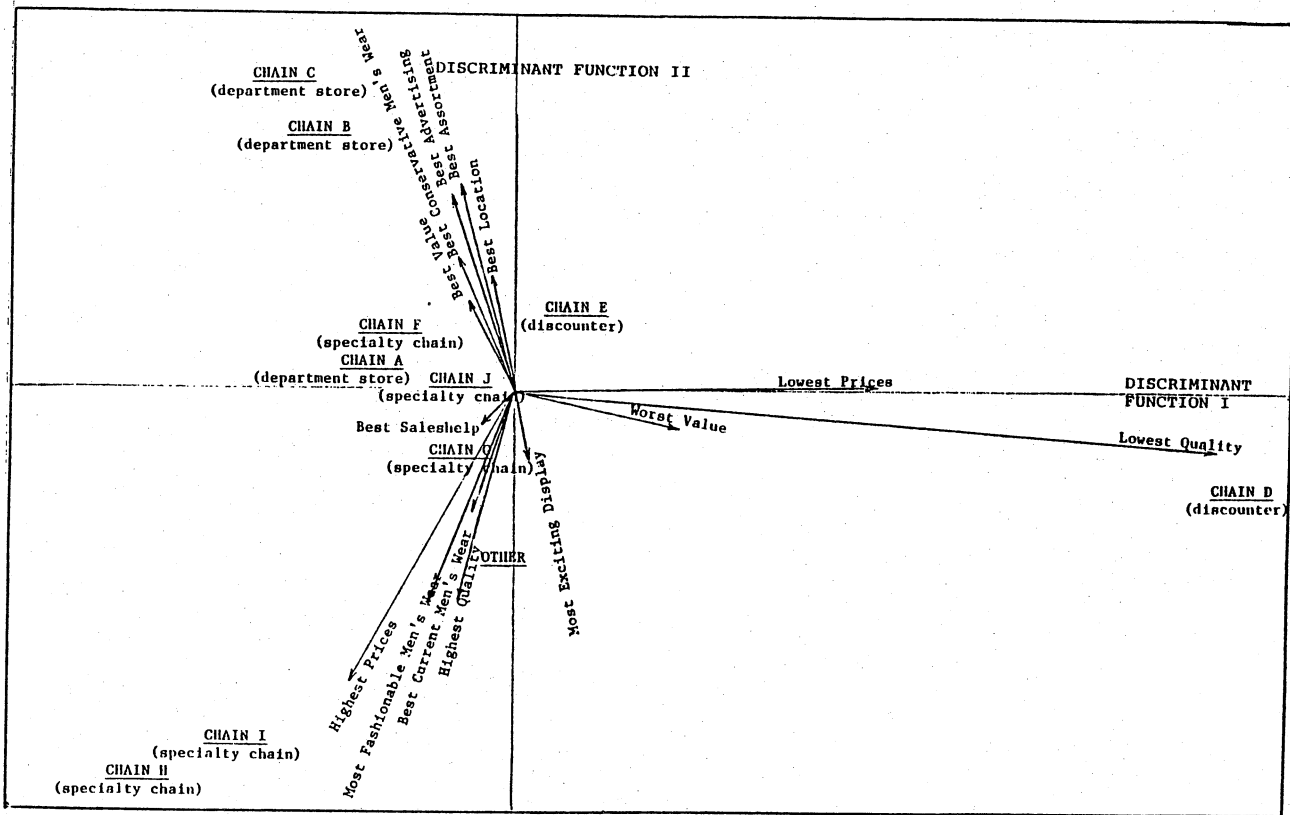


FIGURE 1

Perceptual Map of Toronto Men's Wear Store Space Displaying Store Centroids and Determinants of Patronage Vectors on Discriminant Functions I and II.

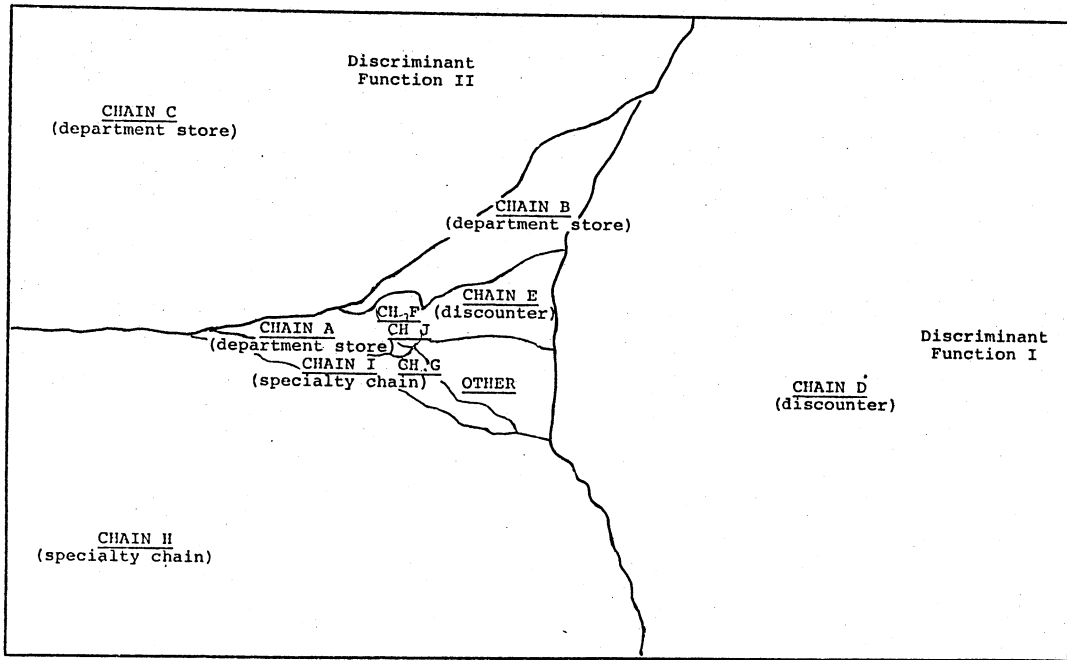


FIGURE 2

Territorial Map of Toronto Men's Wear Market on Discriminant Functions I and II, 1975.

This is not a missing page. The editor got mixed up on the page numbering.

H.K.H.

TABLE 6
Centroids on the First Eight Discriminant Functions for Eleven Toronto Chains

Chain	Group	DFI	DFII	DFIII	DFIV	DFV	DFVI	DFVII	DFVIII
Chain A	1	-.23988	.07862	.21648	-.30772	.06943	-.03179	-.04393	.03046
Chain B	2	-.34053	.57478	-.29208	-.08575	.00797	.07625	.01902	-.06346
Chain C	3	-.39050	.60686	-.26839	.06753	-.02490	.00072	-.00919	.06910
Chain D	4	2.20117	-.06829	-.14290	-.01948	-.01529	.01260	-.01145	.00639
Chain E	5	.04651	.10295	.26555	-.03737	-.16388	-.05371	.11551	.00019
Chain F	6	-.16458	.10324	.04272	.22061	-.02995	-.09721	-.04186	-.01446
Chain G	7	-.04759	-.03864	.21544	.01545	.01348	.04103	-.03915	-.05518
Chain H	8	-.54141	-.78425	-.41722	-.06998	-.02397	-.08474	.00826	-.00959
Chain I	9	-.38645	-.48617	.07909	.06784	-.06699	.18692	-.00066	.03062
Chain J	10	-.13605	-.00221	.20657	.04839	-.05859	.04181	-.05649	.00180
Other	11	-.00068	-.09130	.09473	.10048	.29269	-.00826	.05994	.00774

READ: The centroid of Chain A on Discriminant Function number 1 was -.23988 and the centroid of Chain A on Discriminant Function number 2 was .07862.

References

- Leonard J. Berry, "The Components of Department Store Image: A Theoretical and Empirical Analysis", Journal of Retailing, 45 (Spring, 1969), 3-20.
- Irving Burstiner, "A Three-Way Mirror: Comparative Images of the Clienteles of Macy's, Bloomingdale's, and Korvette's." Journal of Retailing, 50 (Spring, 1974), 24-37.
- Theodore Clerenger, Jr., Gilbert A. Lazier, and Margaret Clark, "Measurement of Corporate Images by the Semantic Differential." Journal of Marketing Research, 2 (February, 1965), 80-82.
- Robert J. Dolan, "Market Segmentation Via Alternative Discriminant Procedures," in The 1975 Combined Proceedings of the American Marketing Association, ed. by E. M. Mazze, (The American Marketing Association, 1975), 132-136.
- Richard M. Johnson, "Market Segmentation: A Strategic Management Tool." Journal of Marketing Research, 8 (February, 1971), 13-18.
- Charles W. King and Lawrence J. Ring, "The 1975 Toronto Male Fashion Research Project: A Descriptive Overview." Institute for Research in the Behavioral, Economic, and Management Sciences, Purdue University, No. 526. (Lafayette, Indiana: Krannert Graduate School of Industrial Administration, Purdue University, September, 1975).
- Joseph B. Kruskal, "Nonmetric Multidimensional Scaling: A Numerical Method." Psychometrika, 9 (June, 1964), 115-29.
- Eleanor G. May, "Management Applications of Retail Image Research", Marketing Science Institute Working Paper. (The Marketing Science Institute, Cambridge, Massachusetts, 1973).
- Lawrence J. Ring, Douglas J. Tigert, and Charles W. King, "Fashion Adoption, Retail Image, and Lifestyle: An Integrated Research Program", a paper presented before the

83rd Annual Convention of the American Psychological Association, Division 23, Chicago, Illinois, (September, 1975).

Ricardo L. Singson, "Multidimensional Scaling Analysis of Store Image and Shopping Behavior." Journal of Retailing, 50 (Summer, 1975), 38-52.

Douglas J. Tigert, The Changing Structure of Retailing in Europe and North America: Challenges and Opportunities, a monograph. (The University of Toronto Press, January, 1974).

A HIERARCHICAL MODEL OF SOURCE EFFECT
IN RETAIL NEWSPAPER ADVERTISING

James G. Barnes, Memorial University of Newfoundland, Canada

Abstract

The relative source effects of retail store name, brand name and newspaper on consumer reaction to retail newspaper advertising are examined. Significant retail store and brand effects are found. Mean scores on three dependent variables reveal that retail advertisers can utilize different brand strategies in their newspaper advertising in order to obtain more positive consumer reaction.

Introduction

It has been generally accepted that the effectiveness of an advertising message in accomplishing the objectives of the communicator will depend in part on the attitudes of the audience toward the originator of the message and toward various other sources involved in its transmission. The influence of the various sources in affecting consumers' reactions to advertising will depend on a number of attitudes towards these sources. These attitudes relate to feelings of affection, admiration, fear, power, and credibility. Hovland, Janis and Kelley (1953) concluded that such attitudes are the result of a learning process.

Research by marketing theoreticians into the concept of source effect has shown that the reactions of an audience to a communication are generally affected by the audience's perception of the communicator's expertise and trustworthiness in the communications context (Cox, 1963). Aaker and Myers (1975) have also observed that in choosing a particular spokesman to represent a brand, consideration should be given to the credibility (impartiality and expertness), attractiveness (prestige), and power of the source. They add that credibility and attractiveness are the two most appropriate dimensions of source for advertisers.

One failing, however, of much research which has been done in the advertising area on source effect is that studies have generally examined the effect of a single source. A number of studies have examined the effect on a consumer's reaction to a marketing communication of perception which a consumer has of the salesman delivering the message. Jacoby, et al. (1974) have examined the effect of a smiling source. McDougall and Fry (1975) investigated the effect of retail store image on consumer reaction to retail advertisements.

A number of authors have, however, commented that the concept of source effect should not necessarily be viewed in the context of the single source. Hovland, Janis and Kelley (1953) have stated that regardless of whether a source is perceived as the originator of the message, as the channel through which the message passes, or as an endorser who is featured in the message, the same basic principles apply in that the attitudes of members of the audience toward either of these sources may influence the manner in which the message is received and behavior which may result from receipt of that message. It should be clear that, in the advertising context, a number of sources are at work simultaneously in influencing the receipt of a marketing communications message. Fuchs (1964) examined the effects of two components (company and magazine) of what he considered to be a "total source" with several natural components. He went on to speculate that source

is likely "n-dimensional" and that it is a differently-constructed total complex for each individual. This suggests that the way in which an individual perceives the several sources associated with a particular advertisement will depend on his or her past experience with the manufacturer of the product, with the retailer, and with the 'medium vehicle' used to transmit the message. Since the experiences of individuals will vary, the total effect of the complex of sources will likewise vary across individuals.

Hansen (1972) also suggests that the effect of the various source components will be variable. He concludes that the individual's perception of the source will depend partly on the medium which carries the message and partly on the company which initiates it. The relative influence of Hansen's two source components will vary with the consumer's familiarity with each. He suggests that where a nationally-known manufacturer advertises in a large magazine, the source effects may be ascribed mainly to the company but where the advertiser is not well-known and the publication somewhat more dubious, the source effect may be based more on perception of the magazine.

This multi-levelled concept of source effect is largely substantiated in the research of Fuchs (1964). Using advertisements featuring unknown products, he found that subjects' attitudes towards these products were influenced by the company sponsoring the advertisements and by the magazines in which they were carried. He also found that the greatest attitude change towards the product was accomplished by the combination of high-prestige company and high-prestige magazine, and the smallest attitude change toward the product was attributable to the combination of low-prestige company and low-prestige magazine. The attitude towards the product changed as a linear function of the prestige of the two source components.

These research results suggest that in retail advertising where the brand name of the advertised product is featured, consumer perceptions of the advertised product may be influenced by perceptions of the manufacturer. Where the consumer is not familiar with the brand name or where the brand name of the advertised product is not present, perception of the quality and value for the money involved in the purchase offer may be influenced by the consumer's attitude towards the retail store. It may be hypothesized further that only when the consumer is not familiar with either the brand name of the product or the retail store sponsoring the advertisements is the perception of the purchase offer likely to be influenced significantly by the medium vehicle which carries the advertisement.

¹ Fuchs (1964) rated each company and magazine using five semantic differential-like scales: good-bad, reliable-unreliable, trustworthy-untrustworthy, prestigious-not prestigious, and reputable-disreputable. Companies and magazines which scored high on these scales were labeled 'high-prestige' and those which scored low were labeled 'low prestige'. Possibly more appropriate labels would have been 'positive-image' and 'negative-image'.

The Hierarchical Model

It is useful to examine the several sources which may operate in an advertising context, and to distinguish sources of the product from sources of the message. In the case of a national advertiser whose advertising is carried in a popular consumer magazine, the company may be considered the primary source of both the product and the message. The magazine which carries the advertisement may be considered a secondary source of the message. If that advertisement had featured a well-known celebrity endorsing the product, that person could have been considered a tertiary source of the message. In the case of retail advertising where the brand name of the advertised product is supplied, we may regard the manufacturer of the product as the primary source of the product and the retailer as the secondary source of the product. Various middlemen may be involved in the actual physical movement of the product from manufacturer to retailer, but these are not relevant to the formation of consumer perceptions of the advertised product.

Also, in this branded retail advertisement, the retail store may be considered the primary source of the advertisement since it is seen by the consumer as the sponsor. The newspaper which carries the advertisement is a secondary source of the message and any persons featured in the advertisement as endorsers of the featured product should be considered tertiary sources of the message. Others, such as advertising agencies, may be involved in the preparation of the advertising message, but these are again not relevant to the formation of a consumer reaction to the advertised product.

In retail advertising where the brand name of the manufacturer of the product is not present, as is the case with much fashion advertising, the retail store is perceived as the primary source of both the product and the advertised message.

This study is designed to test for the presence of a hierarchical source effect in retail newspaper advertising. Because the use of an endorser in retail advertising is rare, the study concentrates on the relative impact of the manufacturer's brand name, the retail store's name, and the newspaper in influencing consumers' reactions to retail newspaper advertising. Previous research on the existence of a dual source effect (Fuchs, 1964) suggests that different source effects will operate in different contexts. For example, it is possible to test whether a brand name with a positive image is sufficient to offset a negative attitude towards a retail store in influencing a consumer's purchase decision. Similarly, it is possible to determine whether a retail store which was recently established or which enters a new market area can produce positive responses to its advertising by featuring brand name products which consumers hold in high regard. The different contexts in which a retail advertisement might appear are largely controllable by the retail advertiser. While the retailer must, for the most part, operate within the constraints of his existing store image, which may be positive, negative, or non-existent in the case of a new store, he does have some control over the brand names which he chooses to feature in his advertising, and over the newspapers in which the advertising appears. In making the selection of brand name products to appear in his advertising, the retailer may choose to feature brands which have a positive image, brands which are considered to have a negative image in the consumer's mind, or brands with which the consumer may be unfamiliar, or he may decide not to feature a brand name product in his advertising at all. Similarly, in many metropolitan markets, the retailer has a definite choice to make with respect to the newspapers in which his advertising is placed.

The implications for the retail advertiser should be obvious. Since the retailer is able to manipulate at least two of the source variables, the results of this study should permit questions such as the following to be answered:

1. Where a retailer has an established reputation and image in the trading area, is the newspaper selected likely to influence consumers' reactions to an advertisement?
2. Is it possible for an established retailer to detract from the effectiveness of his advertising by featuring in an advertisement a brand name which has a negative image?
3. In a situation where a retailer is moving into a new market and consequently has no established image, is it possible to improve reaction to newspaper advertising by featuring brand names which have a positive image and/or by advertising in a newspaper which has such an image?
4. Can a retailer whose store has an established negative image improve reaction to newspaper advertisements by featuring brand names which have a more positive image?

Hypotheses

In retail newspaper advertising, there are many different possible combinations of brand name, store name, and newspaper. It is, therefore, reasonable to hypothesize that different combinations of these variables will produce different consumer reactions. For example, a manufacturer's brand name may be present or absent; if present, the brand name may be known or unknown; and, if known, it may convey a positive or a negative brand image. Similarly, the name of the retail store appearing in the advertisement may be known or unknown; and, if known, the store image may be positive or negative. Finally, in cities where more than one daily newspaper operates, a selected newspaper may be held in high regard, or may have developed a negative image. A series of hypotheses have been developed based on the research reviewed above and on earlier study conducted by the author (Barnes, 1975).

The hypotheses to be tested in this research are as follows:

1. That the perceived source of the product will have a greater effect on the reaction than will the perceived source of the message;
2. That the brand name of the advertised product will have a significant effect on subjects' reactions;
3. That, in situations where no brand name is present or where the brand name is unknown, the retail store effect will predominate;
4. That the effect of the newspaper in which the advertisement appears will be significant only in situations where the brand name is either absent or unknown and the retail store is unknown.

Research Design

Independent Variables

The three main independent variables whose effect on consumer reactions to retail advertisements is examined

in this study are the three elements of source effect discussed above. The three main treatments are, therefore, manufacturer's brand, retail store, and newspaper. In order to incorporate the various possible combinations of brand, store, and newspaper, the following levels of each of the three main treatment variables are incorporated into the study:

- BRAND NAME:
1. positive brand image
 2. negative brand image
 3. unknown brand name
 4. no brand name present
- RETAIL STORE:
1. positive store image
 2. negative store image
 3. unknown store name
- NEWSPAPER:
1. positive image as source of information
 2. negative image as source of information

The research involves a 4 x 3 x 2 design which necessitated the preparation of twenty-four different versions of a print advertisement, each of which was exposed to subjects in one of the twenty-four cells of the design. In addition to testing the effects of these three main treatment variables, data were gathered on subjects' attitudes, interests and opinions concerning retail shopping and the mass media, attitudes towards specific stores, local newspapers and advertising in general, and certain demographic and socioeconomic information relating to the respondent and to her family.

Dependent Variables

In accordance with a suggestion made by Olson (1974), this study incorporates three dependent variables: perceived believability of the purchase offer; perceived value for the money; and motivation to act. The position of a respondent on each of these variables is considered to represent a component of her total reaction to the advertisement to which she is exposed. These three components of reactions have been included because this provides a more accurate indication of the subject's potential purchase behavior with respect to the advertised product. Insofar as the subject's reaction to the advertised purchase offer may be considered to represent her attitude toward the product in the context of the advertisement, the three-fold representation of that reaction may also be considered to correspond to Rokeach's conceptualization of attitude as "an organization of beliefs" (Rokeach, 1972). Rokeach conceived of each belief in an attitude organization as having three components: cognitive; affective; and behavioral. The dependent variables in this study may be interpreted as corresponding to these three components. The subject's believability of the advertisement may be considered to represent a person's knowledge about what is true or false, good or bad, desirable or undesirable. Perception of value for the money represents an affective dimension of the subject's reaction to the advertised product. The third dependent variable, motivation to act, clearly represents a behavioral dimension.

Questionnaire Design

The questionnaire employed in this study contained four major sections as follows:

1. Statements relating to subjects' activities, interests, and opinions concerning retail shopping and the mass media. Subjects indicated their level of agreement with each statement on a seven-point, Likert-type scale.

2. Questions relating to subjects' attitudes toward specific retail stores in the trading area, toward local newspapers, and toward advertising in general.
3. A retail newspaper advertisement to which subjects were asked to express their reactions in terms of the three dependent variables described above.
4. Demographic and socioeconomic questions relating to the respondent and to her family.

The three criterion variables (perceived believability, perceived value for the money, and motivation to act) were operationalized in the third part of the questionnaire. This part of the questionnaire contained twenty-one statements, each of which forms part of one of three scales, representing believability, value for the money and motivation to act. Subjects were asked to respond to each of these twenty-one statements on a seven-point, Likert-type scale ranging from definitely agree to definitely disagree.

The advertisement which appeared in the third part of the questionnaire featured sheets and pillow cases, as products with which female consumers would be familiar. The actual advertisement was designed by the advertising department of a local newspaper in such a way as not to reflect the advertising style of any particular retailer. The advertisement was prepared in twenty-four different versions, with each version featuring the name of a different retail store, a brand name for the advertised products, and the folio of a local newspaper.

The retail stores, brand names and newspapers featured in the advertisements were selected and labeled 'positive image' and 'negative image' based on research previously conducted in the area. Earlier store image studies were used to identify a local department store which engages in very little sale advertising and which has a very positive image in the market. This store was labeled 'positive image'. A discount store which employs a strong promotional advertising approach and which has a rather negative image was labeled 'negative image'. The assignment of these stores to their respective levels of the retail store variable was substantiated by the store image data gathered as part of this survey. A fictitious store name was used in one-third of the advertisements to represent the 'unknown' store level of the retail store variable.

A preliminary survey among female heads of households obtained data on the image of various brands of sheets and pillow cases. A brand which demonstrated high awareness and a high quality image was labeled 'positive image' and a second brand also demonstrating high awareness but a low quality image was labeled 'negative image'. A third fictitious brand name was used to represent the 'unknown brand' level of the brand name variable. One-quarter of the advertisements contained no brand name.

Earlier research studies on newspaper readership and image were used to assign the two local daily newspapers to the 'positive image' and 'negative image' levels of the newspaper variable. This assignment of the newspapers was also substantiated by data collected in this survey. Each of the twenty-four versions of the advertisement employed the same copy and prices.

Data Collection

The sample was drawn from the retail trading area of an Eastern Canadian city. This city was selected because it offers a broad range of large department and discount stores and because it has two daily newspapers. A total of 2,400 questionnaires was mailed to female heads

of households selected randomly from recent voters' lists. Completed questionnaires numbered 429, of which 402 were useable. One hundred copies of each version of the advertisement were mailed and the 402 useable questionnaires returned resulted in cell sizes ranging from fifteen to twenty-five.

TABLE 2

FACTORS PRODUCED FROM PRINCIPAL COMPONENTS ANALYSIS OF 'LIFE STYLE' STATEMENTS (n=402)

Analysis

Subjects' responses to the twenty-one statements relating to the advertisement were analysed using principal components analysis (Nie, 1975) and three factors resulted which were clearly interpretable as the three criterion variables discussed above. Sum scores on these factors constituted subjects' scores on the dependent variables. Results of this principal components analysis are presented in Table 1.

TABLE 1

FACTORS PRODUCED FROM PRINCIPAL COMPONENTS ANALYSIS OF 'ADVERTISEMENT' STATEMENTS (n=402)

	FACTOR LOADINGS	INTERNAL CONSISTENCY COEFFICIENT
FACTOR 1: Eigenvalue = 6.14907 - "motivation to act"		
This advertisement would catch my attention if I saw it in a newspaper	0.74287	0.7474
I would likely remember this advertisement.	0.71260	0.7503
This advertisement would attract many shoppers.	0.69750	0.6696
This advertisement certainly interests me.	0.66391	0.7593
I should probably buy these products now while I have the chance.	0.63864	0.7766
If I were shopping in this store today, I would probably take a look at these products.	0.61276	0.7354
Most of my friends would be interested in buying these products.	0.56120	0.7003
FACTOR 2: Eigenvalue = 1.92602 - "perceived value for the money"		
This is probably the best price at which I can buy these products in St. John's.	0.75704	0.7945
If I were to buy these products at another store, I would have to pay a higher price.	0.61662	0.6930
This looks like a pretty good deal to me.	0.55301	0.7829
These products at these prices represent good value for the money.	0.49450	0.6909
I can get a better deal than this at another store.	-0.47813	0.5860
I can get better value than this at another store.	-0.47127	0.6657
FACTOR 3: Eigenvalue = 0.89249 - "perceived believability"		
I don't believe ads like this.	0.68636	0.7257
This advertising is misleading.	0.65401	0.6657
The quality of these products is probably low.	0.59954	0.6540
I don't trust this store.	0.55621	0.6661
This is probably old merchandise that the store is trying to get rid of.	0.52398	0.5846
This advertisement is honest.	-0.42853	0.5697

The sixty "life style" statements in the first part of the questionnaire were also subjected to principal components analysis and eight factors were obtained with eigenvalues greater than unity. These eight factors are interpretable as dimensions of consumer behavior relating to retail shopping and the mass media and are important insofar as they may influence consumer reaction to advertising. The results of this principal components analysis are presented in Table 2. These eight "life style" factors and a number of demographic variables were included in the analysis as covariates since it is hypothesized that reaction to an advertisement is not a function only of the brand name, store name and newspaper associated with that advertisement, but also of certain demographic and life style characteristics of the respondent herself.

	FACTOR LOADINGS	INTERNAL CONSISTENCY COEFFICIENT
FACTOR 1: Eigenvalue = 6.16498 - "bargain orientation"		
I shop a lot for "specials"	0.78257	0.7576
I read newspaper advertisements closely, looking for bargains.	0.68378	0.7252
I consider myself a real "bargain hunter"	0.61006	0.7289
Each week I look in the newspaper for the grocery specials at the supermarket	0.60353	0.5855
It is not worth the extra time and effort it takes to shop around for bargains	-0.52881	0.6473
A person can save a lot of money shopping around for bargains	0.50125	0.6462
I like shopping at "sales"	0.48102	0.6429
FACTOR 2: Eigenvalue = 4.22661 - "quality, department store orientation"		
When I shop at a department store, I know that I get good value	0.56524	0.5881
If some brands sell really well, they must be better quality	0.48738	0.5175
High-priced products are always better quality	0.47965	0.5420
Brands which are heavily advertised are always better quality	0.46894	0.4867
When you buy something on sale, you usually don't get a good deal	0.46797	0.5201
You only get what you pay for when you shop	0.44432	0.4859
I would never buy a dress at a discount store	0.44345	0.5509
I would never buy a dress on sale	0.43583	0.4691
I do most of my shopping in department stores	0.42792	0.5924
I feel that I can trust department stores to give me a good deal	0.40372	0.5178
FACTOR 3: Eigenvalue = 3.09405 - "enjoyment of shopping"		
To me, shopping is fun	0.78390	0.7393
I love to go shopping whenever I can find the time	0.70737	0.7346
I enjoy shopping more than more women I know	0.66697	0.6511
I only go shopping when I really need something	-0.62086	0.6070
I like to browse in stores with no specific item in mind to buy	0.61971	0.6869
I love to shop for clothing	0.61920	0.6646
Shopping gives me no satisfaction	-0.48520	0.5084
I enjoy spending money	0.46676	0.5382
FACTOR 4: Eigenvalue = 2.49247 - "reliance on newspapers"		
I don't read newspapers very often	-0.81280	0.6939
I read the newspaper every day	0.76087	0.7614
I rely on newspapers to keep me up to date on what is going on in the world	0.68630	0.7462
When I read a newspaper, I read it thoroughly	0.56597	0.6869
I simply couldn't do without newspapers	0.55756	0.7077
I get more information from reading a newspaper than from watching the news on television	0.33619	0.5811
FACTOR 5: Eigenvalue = 1.99208 - "attitude toward advertising"		
Advertising is usually misleading	0.60367	0.6073
Advertising is dishonest	0.58168	0.6031
The government should prohibit advertising on radio and television in Canada	0.41199	0.6065
I generally believe what I read in department store advertisements	-0.40994	0.5199
There is too much advertising today	0.37298	0.6157
FACTOR 6: Eigenvalue = 1.54094 - "attitude toward discount stores"		
People like me usually shop at discount stores	0.65527	0.7413
I enjoy shopping at discount stores	0.64935	0.7826
When I shop at a discount store, I get good value for my money	0.62660	0.6693
I feel comfortable shopping at discount stores	0.61468	0.7718
FACTOR 7: Eigenvalue = 1.26240 - "attitude toward credit"		
I buy many things with a credit card	0.68722	0.7095
It is good to have charge accounts	0.64034	0.7236
I think every shopper should have a credit card	0.58392	0.6185
I like to pay cash when I buy something	-0.44443	0.5041
To buy anything, other than a house or car, on credit is unwise.	-0.42551	0.6086

TABLE 2 Continued

FACTOR 8: Eigenvalue = 1.00904 - "attitude toward television advertising"		
I like television advertising because it keeps me informed on new products.	0.54111	0.7772
I think television commercials are just as enjoyable as the programs	0.51105	0.6779
Television commercials are accurate in describing the product for sale	0.40179	0.6557

The data were then submitted to the BMD-12V Multivariate Analysis of Variance and Covariance program (Dixon, 1973). Since the twenty-four versions of the advertisement were distributed randomly to subjects, there was no control over the size of cells into which the returned questionnaires would fall. In order to comply with the constraint of the BMD-12V program regarding equal cell sizes, it was necessary to select subjects at random to be eliminated from those cells containing more than fifteen subjects. In all, forty-two subjects were dropped, thereby creating a sample of 360 subjects, equally distributed over the twenty-four cells of the design.

It has been possible in this study, to control the exposure of subjects to the three main independent variables. However, it is also hypothesized that a subject's reaction to an advertisement is influenced by certain demographic and life style characteristics of the subject herself. While it is not possible to select and group together subjects with similar individual characteristics, it is possible, through use of analysis of covariance (ANCOV) to determine the effect on the dependent variables of subjects' positions on these individual characteristics.

The multivariate version (MANCOV) is a useful technique in situations where more than one dependent variable is involved and the researcher is interested in determining the effect of the various independent variables on the total set of dependent measures. In this study, while the effect of the independent variables on each of the dependent measures is of considerable interest, so too is the effect of these independent variables on the subject's overall reaction to a retail advertisement.

A series of analyses of covariance were performed containing a total of sixteen covariates: eight 'life style' variables and eight demographic variables.

Results

Table 3 contains the results of the analyses of covariance performed. In the first three columns of Table 3 are reported the results of the univariate analyses of covariance performed with each of the three elements of subject's reaction as the dependent variables. The fourth column contains the results of the multivariate analysis of covariance which indicates the effect of various independent variables on subjects' overall reaction to the advertisement. The univariate ANCOV, while producing indications of effect on subjects' reactions of the retail store, brand and newspaper variables, do not take into consideration the inter-correlations among the three criterion variables and do not consider whether there are significant differences between the levels of the independent variables in terms of overall reaction to the advertisements, that is, over the set of three criterion variables.

TABLE 3

SUMMARY OF UNIVARIATE AND MULTIVARIATE ANALYSES OF COVARIANCE (n=360)

Source	df	F - Statistics			MANCOV ^e
		(1) Perceived Believability ANCOV	(2) Perceived Value for Money ANCOV	(3) Motivation to Act ANCOV	
Retail Store (S)	2	3.69 ^c	0.67 ^d	0.55 ^d	2.62 ^c
Brand (B)	3	2.07 ^d	2.38 ^d	2.39 ^d	1.62 ^d
Newspaper (N)	1	0.02	0.35	0.07	0.14
S x B Interaction	6	0.62	1.05	2.09 ^d	0.99
S x N Interaction	2	0.05	0.09	0.54	0.45
B x N Interaction	3	0.01	0.33	0.83 ^d	0.52
S x B x N Interaction	6	1.60	0.72	1.97 ^d	1.55
Covariates					
Bargain Orientation	1	0.11	0.86	3.26 ^d	4.07 ^b
Department Stores-Sales Skepticism	1	0.06 ^d	5.91 ^c	6.20 ^c	3.06 ^c
Enjoyment of Shopping	1	2.94 ^d	0.68	4.30 ^c	2.49 ^d
Reliance on Newspapers	1	0.25	0.08	0.10	0.10 ^a
Attitude toward Advertising	1	21.88 ^a	0.50	0.02	8.02 ^a
Discount Stores Orientation	1	0.74	0.75	0.10	0.94
Attitude toward Credit	1	1.56	1.04	0.00	0.89
Attitude toward TV Advertising	1	1.65	0.43	0.30	1.05
Number in Household	1	0.15	0.11	0.00	0.19
Job Outside Home	1	0.09	0.02	0.57	0.44
Occupation	1	0.08	0.26	0.01	0.11
Education Level	1	0.33	0.12	0.22	0.28
Spouse's Education Level	1	0.76	0.00	0.08	0.34
Spouse's Occupation	1	0.45	0.36	0.00	0.60
Age	1	0.36	2.44	1.39	0.83
Family Annual Income	1	0.71	0.44	0.00	0.40
Error	320				

^aSignificant at the 0.001 level

^bSignificant at the 0.01 level

^cSignificant at the 0.05 level

^dSignificant at the 0.10 level

^eThe BMD-12V MANCOV program produces an approximation of the F-statistic

The results of the four analyses indicate that no significant differences are evident between subjects exposed to the two levels of the newspaper variable. The newspaper in which the advertisements appeared had no significant effect on subjects' reaction to the advertisements. The other two main effects did, however, produce significant results.

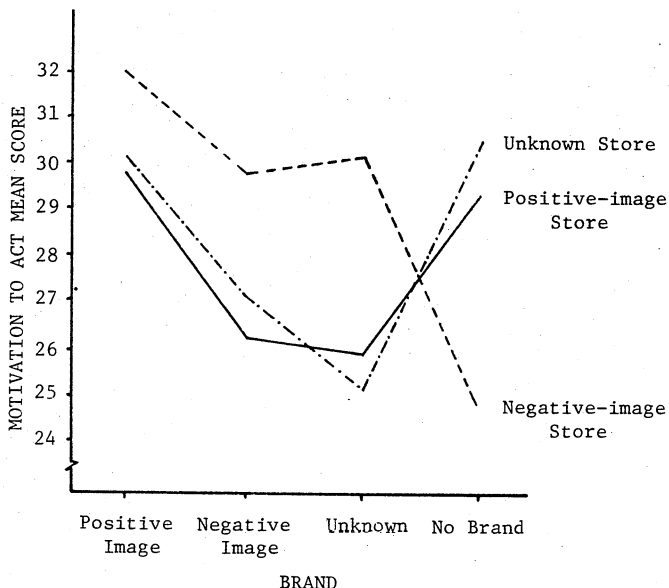
Subjects' perceived believability scores were significantly affected by the retail store sponsoring the advertisement ($p < 0.05$) and by the brand name featured in the advertisement ($p < 0.10$). Two covariates were also significant, suggesting that the extent to which subjects believed the content of the advertisement was influenced by their general attitude toward advertising ($p < 0.001$) and by the extent to which they enjoy shopping ($p < 0.10$).

Perception of value for the money was significantly affected by only one main effect. The brand name appearing in the advertisement was significant at the 0.10 level. Only one covariate was significantly related to perception of value - subjects' orientation toward department stores ($p < 0.10$).

The extent to which subjects were motivated to act was also significantly affected by the brand name ($p < 0.10$). Three covariates were significant in this analysis. Motivation to act scores were significantly related to bargain orientation ($p < 0.10$), orientation toward department stores ($p < 0.05$), and enjoyment of shopping ($p < 0.05$). This analysis also produced a significant two-way interaction (store x brand) effect. Examination of the motivation to act mean scores reveals that the retail store with a positive image obtained its highest scores where it featured a positive-image brand name or where no brand name appeared. Both the negative-image retailer and the unknown retail store obtained high motivation to act scores from advertisements featuring positive-image brands. The ability of the store with a negative image to motivate subjects dropped off somewhat where a negative-image brand or an unknown brand was featured and declined dramatically where no brand name was present. This interaction effect is depicted in Figure 1.

FIGURE 1

TWO-WAY (S x B) INTERACTION
MOTIVATION TO ACT
ANALYSIS OF COVARIANCE



The multivariate analysis (MANCOV) indicates that subjects' overall reaction to the advertisement was significantly influenced by the store sponsoring the advertisement ($p < 0.05$) and by the brand name featured in the advertisement ($p < 0.10$). All four of the covariates which were significantly related to overall reaction were 'life style' variables: attitude toward advertising ($p < 0.001$); bargain orientation ($p < 0.01$); orientation toward department stores ($p < 0.05$); enjoyment of shopping ($p < 0.10$).

Although no significant newspaper effect was found in any of the analyses, mean scores on the three dependent variables by levels of the newspaper variable were examined primarily to explore the possibility that a newspaper effect may be operating where the brand and store effects are absent (that is, where the retail store is unknown and the brand is either unknown or absent). This comparison of mean scores for the two levels of the newspaper variable produced no significant t-statistics. Even where the retail store was unknown and the brand was unknown or absent there were no significant differences between the mean scores produced by the positive-image newspaper and those produced by the negative-image newspaper.

Much information can be obtained by examining mean scores on the dependent variable across combinations of store and brand. These results are presented in Table 4. In the case of the positive-image retail store, while t-tests failed to produce any significant differences among the mean scores of the four levels of the brand variable, there is a definite pattern in evidence. For example, the positive-image store consistently, over all three dependent variables, produced its best scores where its advertising featured either a brand name with a positive-image or no brand name at all. Clearly, in the latter case, the store is able to trade on its established positive store image. Reaction to its advertisements was less positive in those situations where it featured either a negative-image brand or a brand name with which respondents were unfamiliar.

The advertisements for the retail store with a negative image produced even more significant results. In general, this store's advertising was considered less believable than that of the positive-image store. Its most believable advertisements were those which featured the positive-image brand. In terms of indicating good value for the money, the negative-image retail store's advertising produced a significantly more positive reaction where it featured a brand name with an established positive image than where it featured no brand at all ($p < 0.05$). Similar results were found where motivation to act was the dependent variable. In this latter case, a significantly better reaction was obtained even where the negative-image brand or an unknown brand was featured as compared with no brand. It would appear that the negative image of the retail store is imputed to the advertised product in that situation where the product is unbranded.

In the case of a retail store concerning which the consumer has no established image, possibly a new store being established in a market area, it was found that its advertisements were perceived as significantly more believable when the positive-image brand was featured as compared with the situation where an unknown brand was present in the advertisement ($p < 0.01$). The same significant result holds true for the other two dependent variables. An advertisement for an unknown store was more likely to motivate a respondent to act even where it featured a brand with an established negative image ($p < 0.10$) or no brand at all ($p < 0.05$) as compared with advertisements which contained an unknown brand.

TABLE 4
 MEAN SCORES ON THREE DEPENDENT VARIABLES
 BY LEVELS OF THE RETAIL STORE AND
 BRAND NAME VARIABLES
 (n=402)

	Perceived Believability ¹	Perceived Value For The Money	Motivation To Act
<u>Positive-image Store</u> (n=134)			
Positive-image brand (n=34)	17.9194	22.8142	29.3686
Negative-image brand (n=30)	18.9740	20.4258	26.1878
Unknown brand (n=35)	19.4623	22.7525	25.8550
No brand (n=35)	17.7281	23.3357	29.3567
<u>Negative-image Store</u> (n=130)			
Positive-image brand (n=36)	19.0392	25.3533 ^a	30.9470 ^a
Negative-image brand (n=30)	20.6053	23.5654	30.1765 ^b
Unknown brand (n=33)	20.1965	23.8795 ^c	29.5780 ^d
No brand (n=31)	19.9093	20.8640	25.0284
<u>Unknown Store</u> (n=138)			
Positive-image brand (n=30)	18.0667 ^e	25.8788 ^f	30.3642 ^g
Negative-image brand (n=40)	20.5969	23.2373	28.6840 ^h
Unknown brand (N=34)	22.5883	22.3114	24.5440 ₁
No brand (n=34)	20.1492	24.1562	29.3506

¹The lower the score, the more believable is the advertisement perceived.

^aResults are significantly higher for positive-image brands than for no brand at 0.05 level.

^bResults are significantly higher for negative-image brands than for no brand at 0.05 level.

^cResults are significantly higher for unknown brands than for no brand at 0.10 level.

^dResults are significantly higher for unknown brands than for no brand at 0.05 level.

^eResults are significantly higher for positive-image brands than for unknown brands at 0.01 level.

^fResults are significantly higher for positive-image brands than for unknown brands at 0.10 level.

^gResults are significantly higher for positive-image brands than for unknown brands at 0.05 level.

^hResults are significantly higher for negative-image brands than for unknown brands at 0.10 level.

¹Results are significantly higher for no brand than for unknown brands at 0.05 level.

Conclusions and Implications

The brand name main effect was found to be significant at the 0.10 level in all four analyses of covariance. The retail store effect was significant at the 0.05 level in the perceived believability ANCOV and in the MANCOV. No significant newspaper effect was found. These results lend support to the hypothesis that in branded retail advertising, the consumer's reaction is most strongly influenced by those sources which are perceived to have responsibility for the product.

The fact that the newspaper effect was not significant in any case suggests that a consumer's reaction to a branded retail advertisement is not likely to be influenced by the newspaper in which that advertisement appears. A retailer would likely not detract from the attractiveness of his advertising by running advertisements in newspapers which have a somewhat negative image. That is not to say, of course, that all newspapers are equal. While the newspaper may not affect reaction to the advertisement, the retailer must still consider other aspects of the newspapers available to him including circulation, market coverage, cost per thousand, etc.

The most important implications of this study are for the retailer. A quality retail store with an established positive store image can detract from the success of its advertising by featuring negative-image brands or brands with which consumers are not familiar. Its best results

will come from advertisements which feature brands which have a positive image or no brand at all. In this latter case, the store clearly benefits from its established quality reputation.

The opposite effect was found in the case of advertising for a lower quality store which is having image problems. This retailer would also be advised to feature established positive-image brand names in his advertising as these continually produced the most positive reaction. In fact, the negative-image retailer should use even negative-image brands or unknown brands rather than feature no brand at all. Where no brand is present, consumers apparently rely on the image of the store to evaluate the advertised product. Because of its negative image, a retail store which is held in low esteem by consumers should not feature unbranded merchandise in its advertising.

Finally, the retailer who is just establishing a store in a new market area, where he has no established store image, would also be advised to feature established positive-image brands in his advertising. Since the consumer has no image of the store, she apparently judges the advertised product on the basis of the brand name. In this study, advertisements for the unknown retail store produced a more positive reaction where no brand name was present than where either a negative-image brand or an unknown brand was present. Therefore, the retail store which is just establishing itself in a market should avoid the use of negative-image or unknown brands in its advertising.

References

David A. Aaker and John G. Myers, Advertising Management (Englewood Cliffs: Prentice-Hall, Inc., 1975).

James G. Barnes, "Factors Influencing Consumer Reaction to Retail Newspaper "Sale" Advertising," 1975 Combined Proceedings (Chicago: American Marketing Association, 1975), 471-477.

Donald F. Cox, "The Measurement of Information Value: A Study in Consumer Decision Making," In W. S. Decker (Ed.), Emerging Concepts in Marketing (Chicago: American Marketing Association, 1963).

W. J. Dixon (Ed.), BMD: Biomedical Computer Programs (Berkeley: University of California Press, 1973).

Douglas A. Fuchs, "Two Source Effects in Magazine Advertising," Journal of Marketing Research, 1(August, 1964), 59-62.

Flemming Hansen, Consumer Choice Behavior: A Cognitive Theory (New York: The Free Press, 1972).

C. Hovland, I. Janis and H. Kelley, Communications and Persuasion (New Haven: Yale University Press, 1953).

Jacob Jacoby, Edward W. Hart, Jr., George J. Szybillo, and Jacqueline Busato-Schach, "The Persuasive Impact of a Smiling Source," Working paper #138, Department of Consumer Psychology, Purdue University, 1974.

G. H. G. McDougall and J. N. Fry, "Source and Message Content Credibility in Retail Advertisements," Journal of the Academy of Marketing Science, 3(Winter, 1975), 60-68.

Norman H. Nie, et al., Statistical Package for the Social Sciences, 2nd ed. (New York: McGraw-Hill Book Company, 1975).

Jerry C. Olson, "Cue Properties of Price: Literature Review and Theoretical Considerations," Working paper, Pennsylvania State University, 1974.

Milton Rokeach, "The Nature of Attitudes," In Joel B. Cohen (Ed.). Behavioral Science Foundations of Consumer Behavior. (New York: Free Press, 1973).

AN EXAMINATION OF CONSUMER GROCERY STORE CHOICE:
CONSIDERING THE ATTRACTION OF SIZE AND THE FRICTION OF TRAVEL TIME

Daniel J. Brown, Oregon State University

Abstract

Location analysis, commonly known through the gravity model, suggests that consumers balance size against travel costs in choosing retail centers. In contrast to previous research, this paper reports a test of location theory using: individual consumers, individual grocery stores, and consumer panel data instead of survey reports of behavior.

Introduction

In the early decades of this century, Reilly (1931) formulated a retail "gravity model" to predict patronage decisions of rural consumers on the basis of distances to shopping towns and the sizes of those towns. In his model the proportion of trade from an area between two towns, a and b, is defined as B_a/B_b , which is a function of the product of the ratio of the sizes P_a/P_b of the centers (population) multiplied by the inverse ratio of the distances (D_b/D_a) from the intermediate place to the two towns.

$$B_a/B_b = (P_a/P_b) (D_b/D_a)$$

In the early 1960's, faced with a more suburban society, Huff (1962) introduced another "gravity model" to predict consumer choices among shopping centers. In this model, the dependent variable, P_{ij} , represents the probability that a consumer located at i will shop at center j . Predictor variables include, T_{ij} , the travel time required to get from the consumer's travel base to a shopping center, and S_j , the size (floor space) of the center.

$$P_{ij} = \frac{S_j/T_{ij}^\lambda}{\sum_{j=1}^n S_j/T_{ij}^\lambda}$$

Gravity models are still being explored in recent research (Stanley and Sewall, 1976; Bucklin, 1971; Nakanishi and Cooper, 1974).

As models of consumer behavior, they are unusual because of their simplicity. They are based on just two independent variables: the attraction of size and the friction of distance. They also embody easy to visualize, "all other things equal," relationships among variables: a negative monotonic relationship between patronage and travel time, and a positive, monotonic relationship between patronage and size.

The models are related to a theory of consumer choice which is common in the location analysis literature: the consumer weighs the benefit of size against the cost of travel in choosing retail centers. Travel is negatively evaluated because it represents costs to the consumer in purchasing goods (Bender, 1964; Downs, 1961; Kelley, 1958). These costs are measured in terms of

time, out of pocket costs and even emotional energy. On the other hand, size is evaluated positively by the consumer. The underlying attraction might be merchandise assortment as Huff (1962) suggests. In addition, there is a possibility that consumers perceive center size and merchandise prices to be inversely related.

Going beyond the basic gravity model formulation, it is recognized that the attraction of size and the friction of travel will change under different circumstances. Huff (1962), for example, showed that the λ in his model, which captures the trade-off between size and travel, takes on different values when consumers shop for different goods.

When minor expenditures are contemplated by the consumer, the preference for nearby stores would be exaggerated while the preference for size would be reduced. In such a situation, travel costs are large in relation to purchase costs and weigh more heavily in decision making.

When major cash expenditures or a number of purchases are contemplated by the consumer, the reverse is true. The preference for size is exaggerated and the preference for nearby stores is reduced. The consumer can average the cost of the trip over more dollars or more items (Christaller, 1966; Downs, 1961; Bender, 1964). In addition, an outlet which does not allow the consumer to minimize travel costs when only one product is being purchased may allow minimization over the entire trip when the trip involves visits to multiple destinations.

Research Plan

This paper reports on a test of some propositions involving consumer reactions to gravity model variables. Unlike most gravity model research, this study concentrates on a common managerial problem: understanding consumer response to the characteristics of a particular store. Most studies, in contrast, examine shopping centers or towns. Also, unlike most other studies, this research uses the individual consumer as a unit of analysis rather than an entire neighborhood. Individual level analysis has been recommended by several authors (Bucklin, 1967; and Mackay, 1970). A final distinction is that this study uses measures of actual behavior as opposed to reports of behavior from survey questionnaires.

Hypotheses

The following general questions are addressed: Do consumers react to travel friction as if it is a cost to be overcome and to size as a store benefit? Do consumers act as if the importance of these costs and benefits changes with the sizes of payoffs sought on different types of trips? Six specific hypotheses are tested.

The first two hypotheses are fundamental in the basic gravity model formulation examined in the context of grocery shopping.

Hypothesis I: Consumers will make more visits to stores nearby their homes than they will make to more distant stores.

Hypothesis II: Consumers will make more visits to larger stores than they will make to smaller stores.

The second pair of hypotheses explore single-and multiple-purpose trips. A single-purpose grocery shopping trip is made by a consumer when the grocery store is the only destination contemplated and when products are purchased only from the grocery store. A multiple-purpose trip is made by a consumer when some other destination is contemplated in addition to the grocery store.

On multiple-purpose trips, one expectation is that the consumer should be willing to incur greater travel costs. Another expectation is that the consumer should desire to concentrate purchases at a single large center rather than traveling to several different places.

Hypothesis III: Consumers will patronize stores further from their homes on multiple-purpose grocery shopping trips than the centers they will patronize on single-purpose trips.

Hypothesis IV: Consumers will patronize larger shopping centers on multiple-purpose grocery shopping trips than the centers they will patronize on single-purpose trips.

A final set of hypotheses explore convenience and major grocery shopping trips. A "convenience" trip is operationally defined to consist of the purchase of four or fewer items, which are not the specialty of the store, at a cost of \$5.00 or less. Other trips were labeled as "major" if the consumer spent eighteen percent or more of her grocery budget for a month. These cut-off points are arbitrary, but they capture the essence of the distinction being made. Of course, a number of trips did not fall into either category.

Hypothesis V: Consumers will patronize stores closer to their homes on convenience grocery shopping trips than the stores they will patronize on major trips.

Hypothesis VI: Consumers will patronize smaller grocery stores on convenience grocery shopping trips than the stores they will patronize on major trips.

Data

The basic data used to test these hypotheses were taken from a panel of 101 consumers living in a number of suburbs northwest of Chicago. The area was characterized by consistently high levels of store availability and accessibility. It also provided an array of shopping alternatives: "mom and pop" stores, convenience stores, small older supermarkets, and new larger supermarkets.

The panel consisted of white, middle class, suburban females who used the automobile to do their shopping. These women made a total of 992 visits to local grocery stores over a period of a month. Their trips provide the dependent measure in the analysis. Independent measures include: store sizes, shopping center sizes, and travel distances to stores.

Data Analysis

Three methods of data analysis were attempted. Ordinary least squares regression analysis was employed to test

Hypotheses I and II. The basic equation to be estimated was simply:

$$y = \alpha + \beta (x)$$

Hypothesis I predicts that β should be negative, and Hypothesis II predicts that β should be positive. The null hypothesis in both cases is that $\beta = 0$.

The independent variable in Hypothesis I, D_i , represents the travel distance separating consumers' homes from patronized stores. Distance is represented by 20 quarter mile intervals ($i = 1, \dots, 20$). A five mile boundary was chosen because only four of the original 992 trips were made to stores beyond that point.

For any aggregate of shops, it is not too extreme to assume an area to be served. No doubt a native of Detroit might some day buy an article in Birmingham, Alabama, or even in Hong Kong, but it is doubtful whether a retailer of interest to central place analysis takes this into consideration. It seems fair to assume that the territorial ambitions of the most aggressive of hucksters has a limit. To facilitate analysis, it will be assumed here that a local area ends abruptly at its edge, and beyond stretches the void (Curry, 1967, p. 221).

The dependent variable in Hypothesis I, $P_i - O_i$, is a combination of two variables. The behavioral variable of interest, P_i , is the proportion of total visits made to a store within a particular distance interval. The geometry of the spatial choice situation requires the inclusion of a structural variable, O_i , to correct for the "pattern of opportunity". O_i is an index showing that the proportion of the total area in the five mile range which falls within a particular interval. Each successive zone radiating outward from the consumer's home has more area and a higher probability of stores being located there. If distance were not a friction variable, then, more trips would be made to more distant stores simply because of availability.

The independent variable in Hypothesis II, S_j , is the size of the grocery store measured in square feet of floor space. Size was represented by 10 intervals containing 5,000 square feet. The dependent variable, P_j , is the proportion of total visits made to a store in a particular size interval.

Two additional methods were used to test Hypotheses III through VI. The first was a simple test of mean differences between types of trips. The analysis was based on the following hypotheses (Hays and Winkler, 1971, p. 428).

$$H_0: \bar{X}_i - \bar{X}_j = 0 \quad H_1: \bar{X}_i - \bar{X}_j > 0$$

where: X = average value
 i = consumers making one type of trip
 j = the same consumers making the other type of trip

The final method for testing Hypotheses III through VI was regression analysis using indicator variables (Neter and Wasserman, 1974, p. 304). An attempt was made to identify separate linear distance decay functions for the different trip types.

Comparative analyses are restricted to those persons who happened to engage in both types of behavior. So the sample sizes vary among tests.

Results

Patronage displays a pattern of "distance decay" envisioned in Hypothesis I, that is, more consumer visits were made to stores in nearby distance intervals than to more distant stores. The pattern is basically negative monotonic as expected. A deviation from the trend appears in the interval from 0 to 1/4 miles. Because of zoning restrictions in the study area, few homes are located this close to stores.

The following regression equation estimates the distance decay function:

$$P_i - O_i = .12765 - .01218 (D_i) \quad r^2 = .89$$

(.001)

Where: P_i = number of visits in distance interval i
 Total number of visits

D_i = travel distance separating consumers' homes from patronized stores

O_i = $\frac{\text{the area in distance interval } i}{\text{total area within 5 miles of consumers' homes}}$

The β coefficient is negative as expected and significantly different from zero ($t = 12.05$, $p < .0005$). Taking the pattern of opportunity into account, almost 90 percent of the variation in the data is explained by distance.

Considering store size in Hypothesis II, the results are not as straightforward. The pattern of patronage is approximately bell shaped over different store size intervals, so linear regression is not appropriate. The pattern is certainly not monotonically increasing as expected. Thus, the data do not indicate that consumers respond to store size as a general benefit.

Consumers who made both multiple-and single-purpose trips acted as expected in Hypotheses III and IV. The results are shown in Table 1. They chose nearer stores when they were buying only groceries and more distant stores when also making non-grocery store purchases. The difference amounts to about a ten percent increase in distance beyond the single purpose stores. However, in regression analysis, it is not possible to isolate two separate linear distance decay functions using indicator variables to represent different trip types. Consumers also chose grocery stores in larger shopping centers when making multiple purpose trips than when making single purpose trips. The difference amounts to about a 96 percent increase in size over single purpose shopping centers.

Consumers who made both major and convenience grocery shopping trips also acted as expected in Hypotheses V and VI. The results are shown in Table 2. They chose more distant stores when the amount of purchase was large and closer stores when purchase amounts were insignificant. The increase in distance is about 27 percent more for major trips than for convenience trips. Again it is not possible to isolate separate linear distance decay functions for the two different trip types. Consumers also chose larger stores on major trips than on convenience trips. The increase of major over convenience amounts to about 40 percent.

TABLE 1

DIFFERENCES BETWEEN MULTIPLE AND SINGLE PURPOSE GROCERY SHOPPING TRIPS: TRAVEL DISTANCES TO AND SIZES OF SHOPPING CENTERS PATRONIZED

	Average Difference	t	d.f.	p <
Travel Distances between Patronized Stores and Homes of Consumers	.12 miles	1.87	77	.05
Shopping Center Sizes Surrounding Patronized Grocery Stores	108,950 sq. ft.	134.1	66	.00

TABLE 2

DIFFERENCES BETWEEN CONVENIENCE AND MAJOR GROCERY SHOPPING TRIPS: TRAVEL DISTANCES TO AND SIZES OF GROCERY STORES PATRONIZED

	Average Difference	t	d.f.	p <
Travel Distances between Patronized Stores and Homes of Consumers	.31 miles	1.88	53	.05
Sizes of Stores Patronized	7,100 sq. ft.	4.60	53	.00

Discussion

The statistical tests employed in this study produce mixed although moderately successful results. Hypotheses I, IV, and VI, are supported completely. Hypotheses III and V receive qualified support. They pass the t test, but the trends envisioned in the hypotheses can not be described by separable linear functions in regression analysis.

Only Hypothesis II receives zero support. "Average" sized stores received more patronage than either larger or smaller stores. The pattern looks as if it could have been generated from a random selection process or as if the most popular stores were average size. Several explanations are possible: First, the results observed may reflect the actual pattern of available stores rather than a preference function for store sizes. Second, the preference function for store size may be non-monotonic as Baumol and Ide (1956) suggest. Size to them is thought to be attractive up to a certain point. Beyond that point increases in size mean more shopping cost for the consumer rather than better selection. Finally, Bucklin's (1967) admonishment about gravity model applications to individual stores may apply:

Doubt that mass will be adequate is brought to

mind because of difficulties experienced by researchers attempting to evaluate inter-store trading areas. At this level of disaggregation, the image content of shopping utility evidently becomes the controlling factor (p.38).

However, this final assertion seems inconsistent with the favorable result for Hypothesis VI which indicates that consumers were reacting to store size as a source of utility.

The results of this study lend credence to the general location theory postulates about consumer behavior. Consumers do appear to react to distance and store size in choosing stores. In addition, the present test is unusual. It demonstrates that location theory can be tested using individual stores and individual consumers as units of analysis. Although Huff (1966) suggests that both of these uses are possible, they are rarely seen. The test is also unique in that it uses actual behavior as a dependent variable rather than behavior reports from surveys.

The study would have certainly been improved by a larger sample size. Then both the multiple-single and the major-convenience distinctions could have been investigated simultaneously. The expectation should be that a single-purpose, convenience trip should encourage patronage at the nearest store of any size.

From a managerial perspective, an important implication lies in the realm of trading areas. Smaller trading areas will be controlled by stores which cater to convenience and single-purpose shoppers, and larger trading areas will be enjoyed by stores which cater to major trip and multiple-purpose shoppers.

What happens if the gasoline supply is restricted in the future? Two consumer options are obvious: a reduction in automobile travel or a switch to alternate modes of travel (Becker, Brown, and Schary, 1976). In the first case, carefully planned multiple-purpose trips or a small number of major trips would be in order for the consumer. The effect could be to increase the trading areas of large stores and large shopping centers.

Switching to non automotive modes of travel (bicycle, walking, and mass transit, for example) generally means that smaller amounts can be brought home from a shopping trip. Furthermore, alternatives such as walking or biking greatly increase the friction of distance. The effect could be to decrease the trading areas of stores so that the economies of scale from large operations would no longer be possible. In the case of either consumer transportation response, the effect on the pattern of retail trade can be expected to be dramatic.

References

W. J. Baumol and W. A. Ide, "Variety in Retailing," Management Science, 3(October, 1956), 93-101.

Boris W. Becker, Daniel J. Brown and Philip B. Schary, "Behavior of Car Owners During the Gasoline Shortage," Traffic Quarterly, 30(July, 1976), 496-483.

Wesley C. Bender, "Consumer Purchase Costs -- Do Retailers Recognize Them?" Journal of Retailing, 40(Spring, 1964), 1-8 and 52.

Louis P. Bucklin, "The Concept of Mass in Interurban Shopping," Journal of Marketing, 31(October, 1967), 38-42.

Louis P. Bucklin, Shopping Patterns in an Urban Area (Berkeley: Institute of Business and Economic Research, University of California, 1967).

Louis P. Bucklin, "Retail Gravity Models and Consumer Choice: A Theoretical and Empirical Critique," Economic Geography, 47(October, 1971), 489-497.

Walter Christaller, The Central Places of Southern Germany (Englewood Cliffs, N. J.: Prentice Hall, 1966).

Leslie Curry, "Central Places in the Random Spatial Economy," Journal of Regional Science, 7(Winter, 1967), 217-237.

Anthony Downs, "A Theory of Consumer Efficiency," Journal of Retailing, 37(Spring, 1961), 6-12.

William L. Hays and Robert L. Winkler, Statistics: Probability, Inference and Decision (New York: Holt, Rinehart, and Winston, Inc., 1971).

David L. Huff, "A Probabilistic Analysis of Consumer Spatial Behavior," Proceedings of the Fall and Winter Conference of the American Marketing Association, 1962, 443-461.

David L. Huff, "A Programmed Solution for Approximating an Optimum Retail Location," Land Economics, 42(August, 1966), 293-303.

Eugene J. Kelley, "The Importance of Convenience in Consumer Purchasing," Journal of Marketing, 23(July, 1958), 32-38.

David B. MacKay, Consumer Movement and Store Location Analysis, Doctoral Dissertation, Northwestern University (Evanston, Ill., 1970).

Masao Nakanishi and Lee G. Cooper, "Parameter Estimation for a Multiplicative Competitive Interaction Model -- Least Squares Approach," Journal of Marketing Research, 11(August, 1974), 303-311.

John Neter and William Wasserman, Applied Linear Statistical Models (Homewood, Illinois: Richard D. Irwin, 1974).

William J. Reilly, The Law of Retail Gravitation (Austin: University of Texas, 1931).

Thomas J. Stanley and Murphy A. Sewall, "Image Inputs to a Probabilistic Model: Predicting Retail Potential," Journal of Marketing, 40(July, 1976), 48-53.

PERCEPTIONS OF UNFAIR MARKETING PRACTICES:
CONSUMERISM IMPLICATIONS

Gerald Zaltman, University of Pittsburgh
Rajendra K. Srivastava, University of Pittsburgh
Rohit Deshpande, University of Pittsburgh

Abstract

Previous research in complaint behavior has ignored the perception of unfair marketing practices as an explanatory variable. Perceptions of unfair marketing practices are related to consumer complaint behavior, although differentially related across different ages. Social involvement among elderly consumers was also found to be a relevant variable. Implications of the data are discussed for public policy makers, consumer educators, and vendors.

Introduction

A critical issue in consumer education is whether perceptions of unfair marketing practices are related to consumer complaint behavior. One reason the issue is important concerns remedial action. If consumers' complaint behaviors are affected by their awareness of unfair marketing practices, then a strategy of consumer education is most appropriate. Increasing consumer knowledge of unfair marketing practices should lead to increased complaint behavior when such practices are encountered. Additionally, educating consumers in the methods of forming and presenting complaints about grievances becomes an important associated issue. These strategies have many advocates (Andreasen, 1975; Kraft, 1977), although at least some persons active in the area question the effectiveness of consumer education programs in increasing consumer knowledge or in altering their behavior (Staelin, 1977; Robertson, et al, 1974; Robertson, 1975). If consumer perceptions of unfair marketing practices are unrelated to their complaint behavior, that is, if barriers exist which prevent consumers from taking action or consumers cannot be instructed effectively, then regulatory action by public agencies may be the most appropriate direction for protecting consumers from unfair marketing practices. This approach essentially says that what consumers cannot or will not do for themselves must be done by regulatory agencies (Palumbo, 1976; Robertson, et al, 1974; Andreasen, 1975, especially Chapters 10 to 12; Feldman, 1976).

This paper is concerned with consumer perceptions of unfair marketing practices and complaint behavior. A special interest is with elderly consumers. The study is exploratory in nature, but has several important implications for consumer educators, public policy makers, and vendors who need to understand consumer complaint behavior and the factors which may lead to such behavior. One particular factor, awareness of unfair practices, which appears not to have been studied to date, is highlighted here. A second factor, social involvement, also unstudied, is mentioned briefly as well. The paper concludes with a discussion of the major implications of the data presented.

Relevant Literature

Consumer complaint behavior has received considerable attention in the literature. Wall, Dickey and Talarzyk (1977) have tried to profile consumer satisfaction and propensity to complain. A very important finding of their research is that dissatisfaction and the experience of a product performance problem was not related to

a consumer's likelihood of complaining. They found that the most effective predictors of propensity to complain were personal characteristics and internal influences. The Wall, et al study did not, however, measure consumer knowledge or awareness of unfair marketing practices. One of their findings on which we will comment later was that "younger consumers were more likely to be noncommunicators of complaints" (p. 33). This finding was not, according to the authors, a particularly strong one.

Another study by Granbois, Summers and Frazier (1977) studied consumer beliefs and expectations as they related to complaining behavior, but did not report particularly strong associations. However, beliefs about a store's willingness to provide a remedy for a complaint was the most important consumer belief variable. By and large, demographic variables such as age were not found to be of significant importance. The Granbois, et al study did not attempt to measure consumer knowledge about practices or events they might complain about.

Landon (1977) reviewed the literature related to consumer complaint behavior to develop a model of consumer complaint behavior. Interestingly, neither the literature he cites nor the model he builds considers the basic issue of whether consumers know what they should complain about.

The most relevant study to the research presented here is reported by Kraft (1977): "...consumers who believe that a firm intentionally deceived them or acted to dissatisfy them will be more likely to complain than those without this perception." Among dissatisfied consumers, those who complained were nearly twice as likely to think sellers deliberately deceived them than consumers who did not complain. This raises the important question of consumer perceptions of unfair marketing practices. Before consumers perceive deception or unfair treatment, they must perceive particular practices to be unfair or deceptive. Kraft's research appears to be the only study which touches upon this important issue.

A special concern we have is with elderly consumers. Waddell (1975) cites several reasons why elderly consumers may be particularly susceptible to fraud. These reasons include lower education than the population at large, increased concern with getting special bargains to make limited resources go further, and greater gullibility and suggestibility arising out of feelings of helplessness, low self-esteem, and loneliness. It has been suggested that the medical area is the area where consumer fraud among the elderly is greatest (Butler, 1975). Other research has focused on information usage of consumers (Klipzel and Sweeney, 1973; Schiffman, 1971) and social isolation (Schiffman, 1973). For example, Schiffman has noted the social involvement in various groups may enhance opportunities for consumption related discussions and hence persons involved in many groups have a greater chance to learn about the market. As noted earlier, Waddell (1975) suggests that the lack of social involvement may be a contributing factor to susceptibility to fraud. Social involvement may provide not only more opportunity to learn about which marketing practices are unfair or deceptive, but may also provide social reinforcement for taking complaining action and the opportunity to learn how or where to complain. Thus, we might expect the socially involved consumer to perceive

more marketing practices as unfair and to be more prone to complain.

Extant literature on consumer complaining focuses almost exclusively on product performance, the related issue of product safety, and to a lesser degree, other postpurchase activities, such as servicing and honoring of warranties (Fisk, 1970; Westbrook, 1977; Darden, 1977; Trombitta and Wilson, 1975). However, consumers may and do complain about many other factors beyond issues related to product performance. It is important therefore to consider a broad range of factors a consumer may complain about. In the next section of this paper, we shall direct our attention to these factors.

Present Research

The data addressed here concern consumer perceptions or recognition of unfair and deceptive practices and their undertaking complaint behavior. As evidenced in the review above, this has been largely unstudied in the literature on consumer complaint behavior. Moreover, this study reports on consumer recognition of unfair practices at three points: the point or stage prior to product purchase; the time of purchase stage; and postpurchase stage. This distinction does not appear to have been made before. Additionally, this study addresses the issue of age as it relates to the recognition of unfair practices and to complaint behavior. Existing evidence about the relationship of age and consumer complaining is contradictory (Wall, et al, 1977; Gaedeke, 1972; Mason and Hines, 1973; Miller, 1970).

The Sample and Data

The data on which this study is based are derived from a nationwide sample of 2,849 (4,000 questionnaires were mailed and 2,849 usable responses, a net response rate of 79%, were obtained). As the data presented here are merely a subset of a larger research focusing on the consumption problems specifically concerning susceptibility/perception/information processing/complaint behavior of the elderly with respect to fraudulent marketing practices, the sample is naturally biased towards a larger representation of the elderly age group (1771/2849 = 62.2%). In comparison with available statistics on the national level*, the following were the characteristics of the sample:

- (1) A high level of education;
- (2) Slightly unbalanced geographical distribution (lower representation in the New England and East-South Central states);
- (3) Lower representation of the black minority group;
- (4) Lower (substantial) representation of single males;
- (5) Slightly higher income level.

The data are the first wave of a three-wave study. Steps are being undertaken to ensure greater representativeness (within the nonelderly and elderly) subsamples in the remaining stages. However, as we will discuss while examining the results, most of the unrepresentative characteristics noted above are not expected to change the implications of the findings.

Variables and Propositions

The basic proposition tested here is the following: As awareness of unfair and deceptive practices increases among consumers, the propensity to complain increases.

Subpropositions may be formulated for each of three stages: (1) as awareness of unfair practices at the prepurchase stage increases among consumers, the propensity to complain increases; (2) as awareness of unfair practices at the purchase stage increases among consumers, the propensity to complain increases; and (3) as awareness of unfair practices at the postpurchase stage increases, the propensity to complain increases. To test the basic proposition and subpropositions, two variables were constructed. One variable is labelled "unfairness awareness" and is simply measured by the number of practices consumers perceive or recognize as unfair. This was constructed by counting the number of practices consumers cited in each basic stage represented in Table 1. Propensity to complain was measured by the number of complaint behaviors displayed in a critical incident of consumer dissatisfaction. The complaint behaviors which could be cited are shown in Table 2.

Table 1. Practices Perceived Unfair

Practices Potentially Perceived as Unfair	% Citing a Practice as Unfair
A. Before you buy	
Being pressured to buy a product or service other than the one advertised or wanted. . . .	72
Teaching salesmen to use hard-sell pressure. . .	58
Form letters which look as if they were addressed to you personally	53
Packaging, size, or feature differences that make it hard to compare prices directly. . . .	48
Telephone surveys (asking questions about your interests or habits before sales talk)	57
Offers of something for nothing.	68
Advertising items not available in store	62
Product demonstrations that show how easy to use, safe, or effective the product is	21
Unclear advertising, claims, or packaging. . . .	48
Free gifts, give-aways, or contests used to sell products or services.	53
B. What and when you buy	
Being pressured or forced to pay for repairs or services that others were actually responsible for.	60
Companies who change names or locations often. .	51
Being required to pay in advance or make a large down payment to get product/service.	44
Going-out-of-business sales.	33
Unexpected repairs found during servicing or inspection on a product you own.	43
Product sales that support the handicapped or children overseas.	31
Being pressured or forced to pay for repairs or services within a limited time or from a particular person/company.	34
Sales parties in a home.	27
Money-back guarantees.	13
Extra charge for sending repairmen out on weekends	44
Sales involving small monthly payments over a long period of time.	32
C. Problems after you buy	
Paid more than necessary, overcharged.	56
Being charged for undelivered product or service	47
Wrong item or service provided	48
Expensive to use, hidden costs	49
Unsafe product or service.	43
Hard to use product.	39
Repair or replacement parts not available or hard to locate	61
Being told you did not properly use or care for the product.	45
Poor or hard to read warranty.	47
Slow service/repair; repairman not polite. . . .	57

MD = 7%

* Source: U.S. Bureau of Census, General Population Characteristics, Final Report PC(1)-B1 United States Summary (U.S. Government Printing Office, 1972), p. 1-276.

Table 2. Complaint Actions Taken

Actions Taken Due to Problems Encountered	%
Stopped payment or refused to pay.	9
Decided not to buy that product/service or deal with that company again.	43
Complained to family or friends.	30
Asked for replacement or refund.	22
Complained to the person who sold me the product or service	41
Complained to the company or store	35
Complained to a consumer agency.	6
Complained to a public agency or my congressman.	5
Complained in a letter to a newspaper or magazine.	3
Considered taking legal action	8
Consulted or hired a lawyer to protect my interests.	3
I took no action at all.	23
Other.	2

The logic underlying these propositions is that the more practices consumers are aware of which are unfair, the more likely they are to recognize their being the target of unfair practices. The greater consumers' recognition of their being the target of unfair practices, the greater their likelihood of complaining. Moreover, since different unfair practices may warrant different complaint activities, we would expect the range of complaint activities to be broader as the array of practices perceived to be unfair broaden.

We also put forth the following ideas or propositions: (1) unfairness awareness decreases as age increases, and (2) complaint behavior decreases as age increases. These ideas were developed on the basis of exploratory research (not reported in the study) which included in-depth personal interviews with: (1) consumer protection/education/advisory agencies; and (2) convenience samples of individual respondents. However, reasoned arguments could easily be developed to support the opposition propositions. Such studies that do exist relevant to these propositions are contradictory, although we might expect elderly consumers to be less able to complain because of difficulty in returning to a store, having fewer resources to seek legal assistance, greater susceptibility to intimidation, and so on. Also, elderly consumers may participate less in the marketplace and thus have fewer complaints on an absolute scale. However, it was not clear that elderly consumers would have proportionately more or fewer complaints. Unfairness awareness might be expected to increase with age because of greater shopping experience. On the other hand, those who are concerned with consumerism issues and who have benefited from consumer education efforts appear to be younger people (Miller, 1970; Hustad and Pessemier, 1973).

There is an observation concerning the concept of age which the reader must keep in mind. Chronological age is usually a proxy variable. It is used in lieu of other variables which cannot be easily observed or measured. In some instances, age is an important behavioral variable itself. Examples would include an age specific drinking requirement for alcoholic beverages, age specific auto license requirements, and age specific retirement requirements. In other instances, age is simply an indicator for physical health, social attitudes and outlook, medicinal needs, experiences in one or another life situations, and so on. Thus, it is important to ask just what is meant when an observation is made that something such as awareness of unfair practices or propensity to complain varies with age? Unless this question is addressed, nothing is really being explained. Generally, little effort has been made in the consumer behavior literature to clarify what age represents in the contexts in which it is used. This makes it especially difficult to generalize about age or to compare studies involving age since the phenomenon age

is a proxy for may vary from study to study or context to context.

An attempt is made here to indicate what we believe age refers to in this study. In the case of unfairness awareness, we believe age to reflect differing degrees of buymanship skills. We do not believe that older consumers forget or become unclear about unfair marketing practices. Rather, the consumerism movement and its manifestations in classrooms and elsewhere has primarily benefited younger consumers because the movement is a relatively recent phenomenon in terms of its current salience and activity level. The consumerism movement was not pronounced when the present elderly consumer group was younger, and it does not appear to have reached them now (Waddell, 1976; Hustad and Pessemier, 1973). Thus, age may represent differential, effective exposure to consumerist activities and possibly an inability to learn new skills.

With regard to the propensity to complain, age may reflect a lower sense of efficacy or a lower expectation that anything positive may result from registering a complaint. Data in the larger study, but not reported here, are supportive of this contention and research by Klipzel and Sweeney (1971) suggests that feelings of efficacy among elderly consumers may indeed influence their behavior as consumers. Also, knowledge about how to complain may be lower among older consumers as may the physical ability to "get out" to complain.

Analysis

Table 3 contains summary data relating unfairness awareness to propensity to complain. Several subsets of data are presented in Table 3. Three subsets refer to three different points in the purchase process: before buying; when buying; and postpurchase. Three subsets refer to three different age groups: 25-34; 35-64; and 65 and over.* One subset refers to the total number of perceived unfair practices in the entire buying process for each age group. Consider, for example, individuals ranking high on unfairness awareness at the prepurchase stage. For the 25-34 age group, approximately 54 percent of those ranking high on unfairness awareness showed a high propensity to complain about their bad buying experience. For the 35-64 age group, this figure drops to approximately 38 percent. Among the 65 and older age group, only approximately 22 percent of those high in unfairness awareness were high on propensity to complain at this stage. This very same pattern holds for the purchase and postpurchase stages. Similarly, the proportion of consumers at each stage who are over 65 and who are low in unfairness awareness is greater than the comparable group among those aged 35-64, who in turn are proportionately greater than the comparable group among those aged 25-34.

The data in Table 3 support the basic proposition that as awareness of unfair and deceptive practices increases among consumers, the propensity to complain increases. Moreover, this is true for each of the three stages and, of course, for all stages combined. Moreover, we believe the implied causal direction is a valid one. As indicated earlier, being aware of unfair practices enables one to recognize them, which in turn gives the consumer a perceived cause or reason to complain, which

* The breakpoints for the age groups were determined on two considerations: (1) we were specifically interested in behavior of the elderly segment of the consumer market, which in most studies is operationally defined as 65 and over; and (2) it has been postulated that there is a stabilization in life-styles at approximately the age of 35, as by that life stage most individuals have decided their plans regarding career, marriage, etc. (personal communication, Larry Wortzel).

Table 3

Unfairness Awareness: Prepurchase Stage									
Age 25-34			Age 35-64			Age 65 and Over			
Propensity to Complain	Low	Medium	High	Low	Medium	High	Low	Medium	High
Low	31.3	27.0	14.0	51.9	33.3	25.3	69.1	47.4	48.2
Medium	40.6	28.3	32.6	29.5	37.1	36.9	24.7	34.8	30.2
High	28.1	44.7	53.5	18.6	29.7	37.8	6.3	17.8	21.6
	100%	100%	100%	100%	100%	100%	100%	100%	100%
	(n=32)	(n=152)	(n=43)	(n=156)	(n=421)	(n=233)	(n=511)	(n=793)	(n=467)
	x ² ns at .01 level			p<.01			p<.01		
Unfairness Awareness: Purchase Stage									
	Low	Medium	High	Low	Medium	High	Low	Medium	High
Low	45.1	20.0	18.0	51.8	31.3	26.4	66.7	48.6	44.7
Medium	27.5	34.8	26.2	33.2	35.1	37.9	25.7	35.0	31.5
High	27.5	45.2	55.7	15.0	33.6	35.7	7.6	16.4	23.9
	100%	100%	100%	100%	100%	100%	100%	100%	100%
	(n=51)	(n=115)	(n=61)	(n=193)	(n=348)	(n=269)	(n=630)	(n=617)	(n=524)
	p<.01			p<.01			p<.01		
Unfairness Awareness: Postpurchase Stage									
	Low	Medium	High	Low	Medium	High	Low	Medium	High
Low	36.4	28.9	20.5	50.5	29.9	28.5	62.8	43.7	49.7
Medium	36.4	31.3	29.1	35.1	36.9	34.9	28.5	36.4	29.2
High	27.3	40.3	50.4	14.4	33.2	36.6	8.7	19.9	21.0
	100%	100%	100%	100%	100%	100%	100%	100%	100%
	(n=33)	(n=67)	(n=127)	(n=208)	(n=241)	(n=361)	(n=758)	(n=428)	(n=585)
	x ² ns at .01 level			p<.01			p<.01		
Unfairness Awareness: All Stages									
	Low	Medium	High	Low	Medium	High	Low	Medium	High
Low	36.8	26.3	21.3	61.1	31.5	27.4	70.6	47.3	46.5
Medium	36.8	30.7	29.8	28.2	37.3	36.6	23.5	36.6	30.3
High	26.3	43.0	48.9	10.7	31.2	50.0	6.0	16.1	23.2
	x ns at .01 level			p<.01			p<.01		
	100%	100%	100%	100%	100%	100%	100%	100%	100%
	(n=19)	(n=114)	(n=94)	(n=131)	(n=343)	(n=336)	(n=520)	(n=664)	(n=587)
	x ² ns at .01 level			p<.01			p<.01		

is a necessary precondition for complaining. Thus, in seeking to understand complaint behavior we must consider consumer awareness or unawareness of unfair practices. This has been largely ignored to date.

It is evident in Table 3 that age is related to unfairness awareness and to complaint behavior. This can be seen in Table 4 and in Table 5.

Table 4. Age and Unfairness Awareness

Unfairness Awareness Across All Stages				
Age	Low	Medium	High	Total
25-34	8.4	50.2	41.4	100% (n = 227)
35-64	16.2	42.3	41.5	100% (n = 810)
65 and older	29.4	37.5	33.1	100% (n = 1771)
p<.01				

Proportionately, more consumers 65 and older are represented in the low unfairness awareness category relative to the other age groups. Also, relative to the other age groups, proportionately fewer consumers aged 65 and older are in the medium and high unfairness awareness category. Overall, the age group 35-64 seems to have a modestly lower unfairness awareness than the 25-34 age group. Space does not allow us to reproduce the contingency tables for all three purchase stages as displayed in Table 3. However, the relationship between age and unfairness awareness is strongest and unambigu-

ous with regard to awareness of unfair practices after purchase. Fifty-six percent of all consumers aged 25-34 were in the high category at this stage compared to 45 percent and 33 percent for the 35-64 and 65 and older age groups, respectively. Fifteen percent of all consumers aged 25-34 were in the low category at this stage compared to 26 percent and 43 percent for the 35-64 and 65 and older age groups, respectively. Thus, it appears fruitful when considering unfairness awareness by age to distinguish among different purchase stages and the possibly unfair practices especially relevant to each stage.

The relationship between age and propensity to complain appears pronounced. The relevant data on this issue are presented in Table 5.

Table 5. Relationship Between Age and Propensity to Complain

Propensity to Complain				
Age	Low	Medium	High	Total
25-34	25.1	30.8	44.1	100% (n = 227)
35-64	34.6	35.6	29.9	100% (n = 810)
65 and older	53.9	30.7	15.5	100% (n = 1771)
p<.01				

As age increases, the propensity to complain clearly decreases. The reader is reminded that the measure of complaint propensity involved a critical incident which the respondent felt to be a particularly bad experience. Overall propensity to complain might be less for each age group. Since a critical incident approach was used, possible differences among age groups in their participation in the marketplace are highly unlikely to account for the observed differences.

It was indicated earlier that social involvement might be an especially relevant factor affecting perceptions of unfair marketing practices and complaint behavior among the elderly. Published findings about consumerism and the elderly are very few and often without specific empirical data. These studies have suggested, however, that social involvement could be relevant to consumer behavior. Individuals with many personal contacts have more opportunity to learn about causes to complain and to learn about procedures for complaining. Also, social contacts may provide needed social and psychological support or reinforcement for the act of complaining.

Space limitations do not permit a full presentation of the data concerning social involvement among the elderly and awareness of unfair practices and complaint behavior. However, some summary observations are possible. Social involvement was measured simply by the number of different organizations respondents participate in frequently or regularly. These include such organizations as church/temple groups, garden groups, senior citizen centers, clubs and fraternal organizations, volunteer work, family gatherings, and so forth. The measures of social involvement were relevant primarily to older persons and hence reference is made here only to older consumers.

Although a significant chi-square ($p < .01$) was obtained between social involvement and awareness of unfair practices, the substantive significance is modest. For example, 30 percent of those low in social involvement were low in unfairness awareness compared to 32 percent of those high in social involvement. Thus, the relationship is clear, but not of a very high order of magnitude. However, it remains debatable as to what level of difference is "substantial." One way of looking at this issue is to focus on the absolute size of the elderly segment. A percentage point in the elderly segment could mean about 200,000 people.

A relationship between social involvement and complaint activity was obtained similar to that characterizing social involvement and unfairness awareness among elderly consumers. Nearly 60 percent of the persons low on social involvement were low on propensity to complain, compared to nearly 50 percent for those high on social involvement. The chi-square statistic was significantly beyond the .01 level. The relationship between social involvement and propensity to complain is also clear, but not of a very high order of magnitude.

Further research is needed for a better understanding of the impact of social involvement on the consumer behavior of older persons. Interviewing and participant observer techniques might be especially appropriate for this task. Social involvement does appear to have an impact on awareness of unfair practices and complaint behavior. With regard to complaint behavior, a degree of spuriousness may be operating. Consumers who are able to be active in social groups by virtue of physical health, access to transportation, and so forth, may also be able to undertake the efforts necessary to complain; factors enabling a person to engage in social activity outside the home also enable a person to easily return to a store to complain. However, social involvement may still provide necessary social reinforcement for complaining. Further research is necessary to determine whether social involvement leads to avoidance of

situations which might have resulted in complaint behavior. Thus, social involvement, through its impact on unfairness awareness, may reduce the incidence of bad experiences (because the consumers know what to avoid), but once bad experiences occur, it may increase the likelihood of complaints being expressed. The data reported here are consonant with this possibility. Socially involved consumers perceive more practices as unfair and they also complain more.

Summary and Implications

The evidence from this study points up to several implications for different bodies of publics involved in the interface between marketing and society. The relevance of determining consumer awareness of unfair commercial practices and the differential activeness in the marketplace (in terms of both the perception of unfairness and the ensuing complaint behavior) among age groups leads us to make the following suggestions:

Implication No. 1. Traditional market segmentation strategy should include a component to determine how active or passive various segments of consumers are in acquiring product/service related information.

Our study has shown, for example, that older consumers are less active in terms of being aware of unfair business practices and in the means of redressing their complaints. This has consequences for consumer educators in that they should investigate which specific consumer segments are so disadvantaged with a view to developing programs to raise the general level of awareness regarding the particular ways in which unfair practices manifest themselves and the form in which to complain. Vendors should be concerned with attempting to understand how the more active consumers perceive unfair commercial practices so as to better understand their own (or their product/service) image in the eyes of those consumers. Regulatory agencies should clarify the legality of various business practices as perceived by consumers with a view to enacting legislation or implementing regulation to ensure that fairness operates in the marketplace to the satisfaction of both vendors and their customers.

Implication No. 2. The perception of unfair business practices by consumer segments should be further clarified in terms of the current legality or illegality of such practices.

Consumer educators should attempt to learn which practices are clearly illegal and which ones are just due to unclear understanding of the product/service usage. Educators can then serve to disseminate knowledge of these distinctions among the aggrieved market segments.

Vendors should try and assist consumers in the use of various products/services so as to avoid the misunderstandings that result in certain practices being perceived as unfair.

Regulatory agencies should monitor consumer complaints to determine whether regulation is required for currently illegal business practices and also whether changes in law are merited by a sufficient number of complaints regarding what was previously considered to be "legal" commercial activity.

Implication No. 3. The perception of unfair business practices by different consumer segments should be understood in the contents of specific product/service categories to which they might most clearly apply.

Regulators should determine not only what practices were seen as being unfair by consumers, but also the types of products or services which were related to these

practices. Specific industries might be more amenable to an unclear distinction between the legal and the illegal, and therefore more open to such perceptions of unfair activity.

Vendors should be interested in ascertaining whether their particular product or service requires more in terms of clarification in consumer use than law currently prescribes.

Implication No. 4. The perception of unfair business practices by different consumer segments should be understood in the contents of the channels of distribution and promotion that specific product/service categories require and to which they might most clearly apply and also in terms of the purchase stage.

Educators should inform consumers about the particular advantages and disadvantages associated with different forms of distributing and promoting products so that more geographically isolated consumers, for example, are not compelled to be dependent on practices they consider unfair. Vendors and public policy makers should also be interested in knowing which types of media and distribution consumers are leary of and the reasons for their apprehension. The perception of unfair practices by consumers should be considered by educators, regulators, and vendors in terms of whether concerns are particularly expressed at the pre-, during, or postpurchase stages.

Implication No. 5. Within different demographic segments the more and the less aware members should be distinguished in terms of awareness of the fairness of business practices and the means of expressing complaints.

Consumer educators should enlist the aid of the more actively aware within particular demographic segments (e.g., the elderly rural poor) to develop a program to increase general awareness.

Vendors should identify the more active consumers and redress their grievances and also ensure that the less active are not being unfairly taken advantage of. Regulators should coordinate the more active into complaint channels to be able to enact new legislation or implement necessary regulation.

Implication No. 6. Programs should be concurrently developed to both increase the awareness of clearly unfair business practices and the awareness of the means of complaining.

Consumer educators should serve also as data banks on the particular people to contact to redress specific complaints and also the best manner that such complaints should take.

Vendors should participate in the design and conduct of educational programs to ensure that their industries are not unfairly maligned and also so that, if necessary, corrective measures are taken promptly.

Regulatory agencies should see that viable consumer complaints can reach them and that aggrieved consumers never feel that they are helpless when taken unfair advantage of in commercial transactions.

Implication No. 7. The social involvement of elderly consumers in society should be enhanced with a view to increasing both their awareness of fair and unfair business practices and the modes of possible redress of complaints.

Efforts should be made to encourage community groups to address consumer issues, thereby deliberately enhancing the positive impact they have on consumer perceptions of unfair practices and on consumer propensity to complain.

Vendors, public officials, and consumer educators might stimulate such activities by participating in such programs. Most importantly, particular efforts should be made to attract to these special meetings those persons who are typically uninvolved for various reasons in community groups.

References

Alan Andreasen, The Disadvantaged Consumer (New York: The Free Press, 1975).

Robert N. Butler, Why Survive: Being Old in America (New York: Harper & Row, 1975).

William R. Darden, "Satisfaction with Repairs Under Warranty and Perceived Importance of Warranties for Appliances," paper presented at the Research Symposium on Consumer Satisfaction/Dissatisfaction and Complaining Behavior, University of Indiana, Bloomington, Indiana, April 20-22, 1977.

Laurence P. Feldman, Consumer Protection: Problems and Prospects (St. Paul, Minnesota: West Publishing Company, 1976), especially Chapter 8.

George Fisk, "Guidelines for Warranty Service After Sale," Journal of Marketing, 34(January, 1970), 63-67.

Ralph M. Gaedeke, "Filing and Disposition of Consumer Complaints: Some Empirical Evidence," Journal of Consumer Affairs, 6(Summer, 1972), 45-56.

Donald Granbois, John O. Summers, and Gary Frazier, "Correlates of Consumer Expectations and Complaining Behavior," paper presented at the Research Symposium on Consumer Satisfaction/Dissatisfaction and Complaining Behavior, University of Indiana, Bloomington, Indiana, April 20-22, 1977.

Thomas P. Hustad and Edgar A. Pessemier, "Will the Real Consumer Activist Please Stand Up: An Examination of Consumer's Opinions About Marketing Practices," Journal of Marketing Research, 10(August, 1973), 319-324.

R. Eugene Klipzel and Timothy W. Sweeney, "The Use of Information Sources by the Aged Consumer," The Gerontologist, (Spring, 1973).

Frederic B. Kraft, "Consumer Perceptions of Business Fraud: A Key Variable in Complaint and Repatronage Behavior," paper presented at the Southwestern Marketing Association meetings, New Orleans, March 23-26, 1977.

E. Laird Landon, Jr., "A Model of Consumer Complaint Behavior," paper presented at the Research Symposium on Consumer Satisfaction/Dissatisfaction and Complaining Behavior, University of Indiana, Bloomington, Indiana, April 20-22, 1977.

Joseph B. Mason and Samuel H. Hines, Jr., "An Explanatory Behavioral and Socio-Economic Profile of Consumer Action about Dissatisfaction with Selected Household Appliances," Journal of Consumer Affairs, 7(Winter, 1973), 121-127.

John A. Miller, "Who is the Discontented Consumer?" in Thomas V. Greer (ed.), Combined Proceedings (Chicago: American Marketing Association, 1970), 486-490.

Edwin Palumbo, "Consumer Protection," in Hsi-Cheng Li (ed.), Readings in Consumerism (Smithfield, Rhode Island: Bryant College, 1976), 161-185.

Leon Robertson, et al., "A Controlled Study of the Effect of Television Messages on Safety Belt Use," American Journal of Public Health, 64(November, 1974), 1071-1080.

Leon Robertson, "Factors Associated with Safety Belt Use in 1974 Starter-Interlock Equipped Cars," Journal of Health and Social Behavior, 16(June, 1975), 173-177.

Leon G. Schiffman, "Sources of Information for the Elderly," Journal of Advertising Research, (October, 1971), 35-37.

Leon G. Schiffman, "Social Interaction Patterns of the Elderly Consumer," Combined 1972 Proceedings of the American Marketing Association (Chicago: American Marketing Association, 1973).

Richard Staelin, "The Effects of Consumer Education on Consumer Product Safety Behavior," in William D. Perreault, Jr. (ed.), Advances in Consumer Research, Vol. IV, 1977.

William L. Trombitta and Timothy L. Wilson, "Foreseeability of Misuse and Abnormal Use of Products by the Consumer," Journal of Marketing, 39(July, 1975), 45-55.

Fred E. Waddell, "Consumer Research and Programs for the Elderly -- The Forgotten Dimension," Journal of Consumer Affairs, 9(1975).

Fred E. Waddell, The Elderly Consumer (Columbia, Maryland: The Human Ecology Center, Antioch College, 1976), 483FF.

Marjorie Wall, Lois E. Dickey, and W. Wayne Talarzyk, "Predicting and Profiling Consumer Satisfaction and Propensity to Complain," paper presented at the Research Symposium on Consumer Satisfaction/Dissatisfaction and Complaining Behavior, University of Indiana, Bloomington, Indiana, April 20-22, 1977.

Robert A. Westbrook, "A Multivariate Analysis of Post Purchase Satisfaction with Major Household Appliances," paper presented at the Research Symposium on Consumer Satisfaction/Dissatisfaction and Complaining Behavior, University of Indiana, Bloomington, Indiana, April 20-22, 1977.

A COMPREHENSIVE STUDY OF CONSUMER SATISFACTION WITH DURABLE PRODUCTS

Stephen B. Ash (student), Indiana University

Abstract

A preliminary analysis of results from the consumer durables section of a broader study directed by Professor Ralph Day of Indiana University is presented. The results reflect wide variations over the 67 durable product classifications in: (1) the sample proportion reporting a purchase; (2) the relative importance of the items to users; and (3) the patterns of reported satisfaction and dissatisfaction.¹

Introduction

The concern of consumer advocates, policy makers, and the consumer public with the apparent general rise in consumer dissatisfaction has led to a call for more and better research focusing on the products and services which seem to be responsible for the increase in the number of dissatisfied consumers. The design and implementation of "effective" consumer protection programs must overcome two basic constraints: insufficient funding and inadequate information. The first limitation suggests that since policy makers have finite resources with which to support their programs, they must attempt to discern the most useful way in which satisfaction/dissatisfaction (CS/D) research should be conducted. The central decision focuses on what kinds of data should be collected to facilitate the design of programs concerned with consumer satisfaction/dissatisfaction. Also, researchers may consider the related issue of aggregation. In assessing consumer dissatisfaction, should the unit of investigation be at the level of the individual, some particular segment of the consuming population or aggregated data on the entire population? Aggregate analysis is governed by data which may provide only very rough approximations of CS/D trends throughout society. Analysis of disaggregated data, while more costly and time-consuming, promises further understanding of the underlying determinants of consumer satisfaction/dissatisfaction. Although virtually all CS/D research to date has relied on individual measures, there have been some efforts to develop overall indexes of consumer satisfaction (Handy, 1976; Handy, 1977; Lingo and Pfaff, 1972; A. Pfaff, 1972; M. Pfaff, 1976). It is not clear to what extent, if any, the index of consumer satisfaction (ICS) facilitates the actual decision-making process of policy makers. For example, at the industry level, there appears to be some reservation about the willingness of managers to rely on ICS as a decision-making tool (Pratt, 1972).

The adequacy of volunteered complaint information as a basis for assessing consumer dissatisfaction has received recent attention (Day and Bodur, 1977). They suggest that complaint letters may be misleading since they tend not to be representative either of the types of problems confronting consumers or of the types of people experiencing consumer problems. No further elaboration on specific shortcomings of volunteered complaint data (for example, "big ticket" bias) is provided here since these issues have been discussed elsewhere (Day and Bodur, 1977).

Policy makers require a workable basis for setting priorities and planning consumer protection programs. They depend on the availability of the kinds of information which will permit them to evaluate, systematically, existing levels of dissatisfaction with products and services. A recent study (Day and Bodur, 1977) has suggested that at least four kinds of information are required for diagnostic purposes:

1. Number of users in the population.
2. Proportion of users expressing dissatisfaction with the item.
3. The degree of importance associated with the consumption of a product or service.
4. How consumers resolve their dissatisfaction with an item.

They contend that these elements may provide policy makers with a basis for assessing the "need for protection" by consumers of particular goods and services.

Purpose of the Study

The primary contribution of this study consists of reporting preliminary results of analysis on durable goods data obtained by means of a survey research instrument which was developed as a part of a larger research project (Day and Bodur, 1977; Day and Landon, 1976). The current findings may be put into perspective if assessed in relation to the underlying objectives of the parent project. The fundamental goal of the project was the design of a research method which would furnish the kinds of information essential for evaluating consumer satisfaction/dissatisfaction over a comprehensive set of products and services. The operational objective was the development of a product-by-complaints matrix based on "better quality data (than volunteered complaints or 'single incident recall' can give us) of a descriptive nature" (Day, 1977). The matrix would be used by policy makers as a tool for diagnosing consumer dissatisfaction.² At least the following issues were of interest to the designer of the research instrument:

- How and from whom should the data be collected?
- How do volunteered data provided by self-selected individuals differ from those obtained from a scientifically-designed probability sample? Which set provides more representative information about consumer problems over the entire population?
- If these data are collected from a probability sample, how free are they from "big ticket" bias and/or overpowering "demand characteristics" of the data collection procedures?

In general, the relevant information required will consist of measures: (1) of the fraction of the population using the item and (2) of the distribution of users over

¹The data for this study were provided by Professor Ralph Day, Indiana University. The author is grateful to Professor Day and to the reviewers for their comments on the original draft of this paper.

²For further details on the parent project, see reports on 1974 study, Day and Landon, 1976; Day and Bodur, 1977.

the satisfaction/dissatisfaction scale for each separate product/service category. This paper examines the possibility that certain "patterns" reflecting different levels of satisfaction/dissatisfaction may exist for consumers over the various kinds of durable goods. These patterns, if confirmed, might increase the effectiveness of consumer protection programs by enabling policy makers to focus their attention on products with the highest rates of dissatisfaction among users.

Research Design

The data for this research were obtained as part of a survey research project which utilized a set of data collection instruments designed for a larger national project currently in preparation. The actual data for the current study were acquired from self-administered questionnaires which were distributed through the "drop-off/pick-up" method to a probability sample of more than 600 dwelling units in one midwestern city during the fall of 1976. The instrument was extensively pre-tested.³

The questionnaires were quite lengthy, probing both consumer satisfaction/dissatisfaction and complaint behavior on an "aided recall" basis over approximately 200 categories of products and services. The preliminary findings which will be reported here relate exclusively to the usage rate/importance/satisfaction portion of the survey instrument which covered 67 categories of durable products. Items which were believed to be purchased and used by consumers in similar ways were grouped together into categories.

The initial task required respondents to indicate whether or not they had purchased any items from the category over a three-year recall period. Those who had made such a purchase were then asked to provide an indication of the relative importance of that category and of the relative extent of satisfaction or dissatisfaction with items contained in the category. The responses available to respondents for each item were as follows:

Purchase (?) and Importance of these products

- N = I have not bought this product (any of these products) within the past three years.
- L = This product is (these products are) of less importance to me or my family than some of the other durables we own.
- H = This product is (these products are) highly important to me or my family.

How satisfied or dissatisfied I am with them

- 1 = I am quite satisfied with this product (these products).
- 2 = I am somewhat satisfied with this product (these products).
- 3 = I am somewhat dissatisfied with this product (these products).
- 4 = I am quite dissatisfied with this product (these products).

The instrument, itself, was designed to furnish more complete information than is available from volunteered complaint data. Also, it sought to avoid or reduce the bias arising from "single incident recall" surveys which force respondents to limit their responses to one particular experience. By providing data on all products the respondent has purchased over a specified period of recall, the instrument provides a basis for assessing the extent of consumer satisfaction/dissatisfaction with items in

³For details on instrument design and pre-testing, see Day and Bodur, 1977.

each separate product/service category. Another advantage is that it enables the researcher to tap the positive as well as the negative elements of consumers' reactions to a comprehensive set of durable goods.

One familiar approach to the study of CS/D has led to operationalizing the measure of dissatisfaction as the discrepancy between performance expectations for the product or service and actual outcomes (Anderson, 1973; Cardozo, 1965; Day, 1975; Day, 1977; Miller, 1976; Olshavsky and Miller, 1973; Summers and Granbois, 1976; Swan and Combs, 1976). Concern about expectations and ultimate performance has raised important methodological issues regarding their measurement. For example, expectations may be operationalized to include: (1) expectations of product performance, (2) expectations of costs and, (3) expectations of the impact on others (Day, 1977). The relevant research question is whether all sources of expectation need to be assessed simultaneously in order to obtain a representative overall measure of the construct. If so, how is each component of overall expectation to be weighted? Unfortunately, the measurement of perceived product performance is also troublesome. For example, one approach has underscored the need to examine the influence of both physical and psychological dimensions of product performance on consumer satisfaction (Swan and Combs, 1976). The rationale for the research instrument employed in this study is that it lets the respondent define what he means by "satisfied" or "dissatisfied." The subject determines for himself what the points on the scale mean. This enables the researcher to evaluate the pattern of self-defined satisfaction/dissatisfaction over respondents for each product category.

The questionnaire also contains the usual demographics and a set of 14 Likert-type attitude/opinion scales designed to elicit subjects' impressions of business, government and marketing. The balance of the paper will be devoted to the preliminary analysis conducted on the category-by-category responses obtained from 119 usable questionnaires.

Measurement

Data on usage rate, importance, and extent of satisfaction/dissatisfaction are shown in percentages. To examine the associations between the satisfaction/dissatisfaction scores and demographic data, correlation analysis was used. Since the satisfaction/dissatisfaction scales probably should be considered ordinal in nature, use of a nonparametric correlation coefficient seems appropriate. The data consisted of a substantial number of cases classified into comparatively few categories. Since they contained a large number of tied ranks, the Kendall (tau-C) correlation coefficient was preferred to Spearman's rank order correlation coefficient (Nie, Hull, Jenkins, Steinbrenner and Bent, 1975).

During the preliminary analysis, attitude and selected demographic variables, which might be considered as interval-scaled data, were regressed, in step-wise fashion, onto the satisfaction scores. Since examination of predicted criterion scores versus residuals suggested that the data were not suitable for multiple regression analysis, the results are not presented here. Other multivariate techniques may be applied as the analysis of the results proceeds.

Preliminary Results

To demonstrate the kind of information contained by the data and its potential for use by policy makers, two sets of results are presented. First, the category-by-category responses denoting purchase, relative importance and satisfaction/dissatisfaction are summarized and briefly.

analyzed. Next the relationships between demographics and satisfaction are explored.

Product Category Analyses

Tables 1 through 4 contain percentage summaries on 67 categories of durable goods which are grouped in four sections of the durable products questionnaire: (1) housing and home furnishings; (2) appliances and personal care equipment; (3) durable items for entertainment, recreation and education; and (4) cars and other transportation durables. All four tables provide information obtained from 119 usable questionnaires. The first column in each table shows the percent of respondents who reported they had purchased the item within the three-year recall period. The next column shows the percentage of purchasers rating the item as relatively important. The rank order of percent rating the item as important is also provided. This is followed by the percentage distribution of satisfaction/dissatisfaction ratings for each category. As stated earlier, a primary objective of the research plan is to provide a basis for classifying items according to the severity of consumer problems they engender. The classification task is simplified if one relies on a combined measure of dissatisfaction for every product category under investigation. This information is covered by the final two columns of Tables 1 through 4 which show the sum of the two dissatisfaction ratings and the rank order of categories by percent dissatisfied.

In Table 1, the percentage of subjects who reported that they had purchased one or more items in any "Housing and Home Furnishings" category varies widely. Only 1.7% of the sample indicated a purchase of "swimming pool" while 88.0% said they had purchased "draperies, linens, etc." Only three other categories reveal more than 50% of respondents indicating purchase within the specified period of recall: "sofas, chairs"; "housewares, tableware, etc."; and "lamps, clocks, etc." The percentage of consumers who indicated that the various items were highly important to them or their families ranges from 35.4% for "electric blankets" to 100.0% for "single family house" and "heating installation." The percent of buyers who rated an item in one of the two dissatisfaction categories varies from 6.8% for "lamps, clocks, etc." to 37.5% for "mobile home."

A key assumption of the research plan is that the amount of dissatisfaction expressed for an item should be brought into perspective by considering it in relation to the proportion of the population using the product. For example, when a very small segment of the population indicates purchase of an item (e.g., 1.7% for "swimming pools"), reported rates of dissatisfaction probably do not have statistical reliability. It may seem intuitively logical to expect higher rates of reported dissatisfaction for widely-owned items than for products which are used by a relatively small segment of the population. The fallacy of this reasoning is exposed if one compares the rates of usage and dissatisfaction for "mobile homes" with those of "draperies, linens, etc." in Table 1. The category "draperies, linens, etc.", with a reported purchase rate of 88.0%, ranks fourteenth out of twenty according to the percentage of dissatisfied users. On the other hand, the greatest rates of dissatisfaction with "Housing and Home Furnishing" items appear to be associated with those purchased by a comparatively small percentage of the entire population. For example, "mobile homes" are ranked first by percent of dissatisfied buyers even though less than 10% of respondents reported purchase of the product. The number two item according to dissatisfaction is "swimming pools," another product characterized by limited ownership. In Table 4, 36.2% of respondents reported purchase of a used car within the recall period. It seems that approximately 41% of the used car owners were dissatisfied with the product. The apparent high rate of dissatisfaction with used cars reported in this

study is corroborated by the amount of dissatisfaction with the item expressed in conventional complaint statistics. In Table 2, the rates of dissatisfaction reported for "Appliances and Personal Care Products" may be compared to those furnished by another study (Mason and Himes, 1973). It is not surprising that higher levels of dissatisfaction were expressed by respondents in the current study. In their research, Mason and Himes (1973) defined a complaint as "any dissatisfaction with an appliance purchased and for which a consumer seeks relief from an outlet in the channel of distribution." In this study, the survey instrument enabled subjects to register dissatisfaction with an item independent of specific complaint behavior. Comparison of the results of the two studies suggests that the differences in percentage reporting dissatisfaction may be ascribed to the number of dissatisfied consumers who do not undertake any formal search for redress. Table 3 reveals a similar pattern of results and will not be discussed here. The preceding examples are meant to suggest that greater understanding of the kinds of problems which consumers experience is possible if the number of dissatisfied users is considered in relation to the total number of users.

Relationship of Satisfaction and Demographics

One premise for this study is that writers of complaint letters are probably not representative of the full range of people experiencing dissatisfaction with products and services (Day and Landon, 1976; Day and Bodur, 1977; Warland, 1975). Here, the unit of interest is any consumer who feels dissatisfied with his purchase of a durable product, not a complainer per se. The central question is: do dissatisfied users of particular items have anything in common? This question implies that it may be useful to investigate relationships between demographic variables and satisfaction scores over a comprehensive set of durable products. Tables 5 through 8 contain correlation coefficients measuring the extent and direction of the relationships between demographics and satisfaction. Since the entire sample consisted of 119 respondents, cell sizes were checked to minimize the possibility of reporting spurious results. This examination revealed that the response distributions for the education and income variables were skewed slightly to the right of normal. All coefficients reported are significant at the .05 level.

Table 5 contains the correlation coefficients for the "Housing and Home Furnishings" section of the questionnaire. All of the demographic variables are significantly correlated with at least three of the durable product categories. Examination of the table shows varying patterns of relationships between demographics and satisfaction. Although none of the coefficients are large, most of the categories exhibit some relationship between one or more demographic variables and the satisfaction scores. The category with the largest number of significant relationships is the "single family house." On the other hand, no significant relationships appear to exist between demographics and satisfaction for "mobile home" owners. The variables with the largest number of significant relationships were marital status, employment status, home ownership and income. The number of significant relationships for each demographic variable varied widely over the groups of durable products shown on Tables 5 to 8. Another study (Day and Bodur, 1977) examined the relationships between demographic variables and satisfaction scores over a broad set of consumer service categories. They reported:

1. weak associations between demographics and responses on satisfaction;
2. widely varying patterns of significant relationships over these categories;

3. situational characteristics (e.g., marital status, home ownership) were more consistently related to satisfaction scores than most personal characteristics (e.g., education and sex).

The results of the present study parallel the earlier findings of Day and Bodur, 1977.

Discussion

The sample statistics reported here are based on data which were obtained in one midwestern city during the fall of 1976. Since a scientifically-determined probability sample was obtained, the results are probably representative of the sample population. However, any generalization of these results to any other population would be inappropriate.

This study is intended to stimulate more research into the underlying determinants of consumer satisfaction/dissatisfaction with durable products. Some fruitful areas for future research on durable goods may focus on:

- (1) the role of product cost as a basis for predicting consumer satisfaction/dissatisfaction with an item;
- (2) the extent to which consumer expectations can be manipulated by manufacturers or government policy makers;
- (3) the identification and measurement of any "carry-over effect" associated with purchase satisfaction, and
- (4) the impact of contemporary advertising on the gap between consumer expectations and perceived performance for widely-used durable products.

Conclusions

This paper has reported some of the preliminary results of analysis on data collected as part of a larger research project still in progress. These findings cover the section of the questionnaire which elicited satisfaction/dissatisfaction scale responses. On the basis of these results, the discriminating power of the survey instrument as a tool for diagnosing consumer dissatisfaction seems to have been satisfactorily demonstrated. The questionnaire was designed to furnish information on rate of purchase, relative importance and satisfaction/dissatisfaction over 67 categories of durable products. The results showed widely varying relationships between the percent of a sample reporting purchase of a product and the percent of users experiencing dissatisfaction with the item. The relationship between demographic variables and satisfaction were generally weak and varied widely over product categories. On a category-by-category basis, however, interesting relationships between demographics and satisfaction scores were revealed.

The results reported here appear to have face validity. It is hoped that they will demonstrate the usefulness of the information conveyed by the data which was obtained by the survey instrument. There is room for much more research into consumer satisfaction/dissatisfaction with durable products.

References

- Rolph E. Anderson, "Consumer Dissatisfaction: The Effect of Disconfirmed Expectancy on Perceived Product Performance," Journal of Marketing Research, February 1973, pp. 38-44.
- Alan R. Andreasen, "A Taxonomy of Consumer Satisfaction/Dissatisfaction Measures," A paper presented at a workshop on the Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction, sponsored by the Marketing Science Institute (MSI) and National Science Foundation (NSF), April 1976, Chicago, Illinois.
- Richard Cardozo, "An Experimental Study of Customer Effort, Expectations and Satisfaction," Journal of Marketing Research, Vol. 2, August 1965, pp. 244-249.
- John A. Czepiel and Larry J. Rosenberg, "The Study of Consumer Satisfaction: Addressing the 'So What' Question," A paper presented at MSI and NSF Workshop, April 1976, Chicago, Illinois.
- Ralph L. Day, "Extending the Concept of Consumer Satisfaction," Advances in Consumer Research, Vol. IV, 1977 (forthcoming).
- Ralph L. Day and Muzaffer Bodur, "A Comprehensive Study of Satisfaction with Consumer Services," in Ralph L. Day, editor, Consumer Satisfaction/Dissatisfaction and Complaining Behavior, Division of Research, Indiana University, 1977, pp. 64-74.
- Ralph L. Day and E. Laird Landon, "Collecting Consumer Complaint Data by Survey Research," in Beverlee B. Anderson (ed.), Advances in Consumer Research, Vol. III, 1976, pp. 263-268.
- Charles R. Handy, "Monitoring Consumer Satisfaction with Food Products," A paper presented at MSI and NSF Workshop, April 1976, Chicago, Illinois.
- Charles R. Handy, "Indexes of Consumer Satisfaction with Food Products: 1974 and 1976 Survey Results," A paper prepared for the Research Symposium on Consumer Satisfaction/Dissatisfaction and Complaining Behavior, Indiana University, April 21-22, 1977.
- Donald J. Hempel, "Consumer Satisfaction with the Home Buying Process: Conceptualization and Measurement," A paper presented at MSI and NSF Workshop, April 1976, Chicago, Illinois.
- James C. Lingo and Martin Pfaff, "The Index of Consumer Satisfaction: Methodology," Proceedings, Association for Consumer Research, 1972, pp. 689-712.
- William J. Lundstrom and Lawrence M. Lamont, "The Development of a Scale to Measure Consumer Discontent," Journal of Marketing Research, Vol. XII, November 1976, pp. 373-381.
- Joseph B. Mason and Samuel H. Himes, Jr., "An Exploratory Behavioral and Socioeconomic Profile of Consumer Action About Dissatisfaction with Selected Household Appliances," Journal of Consumer Affairs, Vol. 7 (Winter 1973), pp. 121-127.
- John A. Miller, "Studying Satisfaction: Modifying Models, Eliciting Expectations, Posing Problems, and Making Meaningful Measurements," A paper presented at MSI and NSF Workshop, April 1976, Chicago, Illinois.
- Normal H. Nie, G. Hadlai Hull, Jean G. Jenkins, Karin Steinbrenner and Dale H. Bent, Statistical Package for the Social Sciences, 2nd ed., New York: McGraw-Hill Book Co., 1975.
- Richard W. Olshavsky and John A. Miller, "Consumer Expectations, Product Performance and Perceived Product Quality," Journal of Marketing Research, Vol. 9, February 1973, pp. 19-21.
- Anita B. Pfaff, "The Index of Consumer Satisfaction," Proceedings, Association for Consumer Research, 1972, pp. 713-737.
- Martin Pfaff, "The Index of Consumer Satisfaction: Measurement Problems and Opportunities," in H. Keith Hunt, (ed.), Advances in Consumer Research, Vol. V, Association for Consumer Research, 1978, forthcoming.

H. B. Pickle and R. Bruce, "Consumerism, Product Satisfaction-Dissatisfaction: An Empirical Investigation," Southern Journal of Business, Vol. 7 (November 1972), pp. 87-100.

John E. Swan and Linda J. Combs, "Product Performance and Consumer Satisfaction: A New Concept," Journal of Marketing, Vol. 40, April 1976, pp. 25-33.

Robert W. Pratt, Jr., "The Index of Consumer Satisfaction and Corporate Marketing Policy," Proceedings, Association for Consumer Research, 1972, pp. 742-745.

Rex H. Warland, Robert O. Herrmann and Jane Willits, "Dissatisfied Consumers: Who Gets Upset and Who Takes What Action," Journal of Consumer Affairs, Winter 1975, pp. 148-163.

John O. Summers and Donald H. Granbois, "Predictive and Normative Expectations in Consumer Dissatisfaction and Complaining Behavior," A paper presented at ACR Conference, October 1976, Atlanta, Georgia.

TABLE 1

SATISFACTION/DISSATISFACTION RATINGS: HOUSING AND HOME FURNISHINGS

Housing and Home Furnishing Categories	Use/Importance			Distribution of Responses				% Dissatisfied	
	Percent Having Made Purchase*	Percent Rating It Important	Rank By Percent Rating It Important	Percentage of Purchasers Checking Satisfied:		Dissatisfied:		Percent of Purchasers Dissatis-	Rank By Percent Dissatis-
				Quite	Somewhat	Somewhat	Quite	fied	fied
Single Family House	25.0	100.0	1	51.6	35.5	9.7	3.2	12.9	12
Condominium or Coop' Unit	6.0	56.7	12	42.9	28.6	14.3	14.3	28.6	3
Mobile Home	5.2	67.3	9	37.5	25.0	25.0	12.5	37.5	1
Vacation Home	4.3	60.5	11	66.7	16.7	0	16.7	16.7	8
Heating Installation	9.5	100.0	1	33.3	50.0	8.3	8.3	16.6	10
Additions to Home	15.7	55.4	13	55.6	27.8	11.1	5.6	16.7	8
Swimming Pool	1.7	50.0	15	0	66.7	33.3	0	33.0	2
"Do It Yourself" Projects	42.4	39.8	18	49.0	41.2	7.8	2.0	9.8	17
Purchase of Land	8.5	69.4	7	63.6	27.3	0	9.1	9.1	18
Draperies, Linens, etc.	88.0	36.9	19	53.9	34.3	7.8	3.9	11.7	14
Electric Blankets, etc.	26.3	35.4	20	59.4	21.9	12.5	6.3	18.8	5
Baby Furniture	5.9	71.2	6	50.0	37.5	0	12.5	12.5	13
Sofas, Chairs	52.5	66.1	10	50.8	34.9	12.7	1.6	14.3	11
Din. Room, Kitchen Furniture	33.1	69.2	8	38.5	43.6	12.8	5.1	17.9	7
Bedroom Furniture	41.0	77.1	3	56.3	35.4	6.3	2.1	8.4	19
Outside Furniture	19.0	41.1	16	43.5	30.4	13.0	13.0	26.0	4
Floor Coverings	41.0	72.9	4	57.1	32.7	6.1	4.1	10.2	16
Housewares, Tableware, etc.	71.2	52.4	14	56.6	32.5	7.2	3.6	10.8	15
Lamps, Clocks, etc.	62.7	40.5	17	48.6	44.6	4.1	2.7	6.8	20
Items not included above	13.6	72.8	5	54.5	27.3	0	18.2	18.2	6

*N = 119

TABLE 2

SATISFACTION/DISSATISFACTION RATINGS: HOME APPLIANCE AND PERSONAL CARE PRODUCTS

Home Appliance and Personal Care Product Categories	Use/Importance			Distribution of Responses				% Dissatisfied	
	Percent Having Made Purchase*	Percent Rating It Important	Rank By Percent Rating It Important	Percentage of Purchasers Checking		Dissatisfied: Somewhat Quite	Percent of Purchasers Dissatis- fied	Rank By Percent Dissatis- fied	
				Satisfied: Quite Somewhat	Dissatisfied: Somewhat Quite				
Refrigerators, Freezers	28.6	25.2	4	51.4	28.6	11.4	8.6	20.0	7
Air Conditioners, etc.	18.8	10.3	8	30.4	69.6	0	0	0	14
Ranges, Ovens, Grills	23.9	17.9	7	44.8	27.6	20.7	6.9	27.6	4
Clothes Washers, Dryers	30.8	24.8	5	77.8	19.4	2.8	0	2.8	13
Vacuum Cleaners, etc.	42.4	28.0	2	48.0	36.0	6.0	10.0	16.0	10
Sewing Machines, etc.	20.3	9.3	9	54.2	45.8	0	0	0	14
Garbage Disposers, etc.	14.4	6.8	12	52.9	29.4	11.8	5.9	17.7	8
Small Kitchen Appliances	71.8	25.6	3	54.1	32.9	8.2	4.7	12.9	11
Electric Razors, etc.	19.8	8.6	11	50.0	29.2	12.5	8.3	20.8	6
Elec. Hairdryers, Curlers	62.4	22.2	6	43.8	27.4	16.4	12.3	28.7	3
Exercise Machines, etc.	9.4	1.7	14	41.7	33.3	8.3	16.7	25.0	5
Vibrators, Massagers	7.8	1.7	14	20.0	50.0	30.0	0	30.0	2
Eyeglasses, Contact	64.1	59.0	1	61.3	21.3	10.7	6.7	17.4	9
Wearing Aids	1.7	1.7	14	33.3	66.7	0	0	0	14
Medical Appliances, Devices	6.0	3.4	13	50.0	37.5	12.5	0	12.5	12
Items Not Included Above	9.3	9.3	9	28.6	0	0	71.4	71.4	1

*N = 119

TABLE 3

SATISFACTION/DISSATISFACTION RATINGS: DURABLE ITEMS FOR ENTERTAINMENT, RECREATION, AND EDUCATION

Entertainment, Recreation and Education Categories	Use/Importance			Distribution of Responses				% Dissatisfied	
	Percent Having Made Purchase*	Percent Rating It Important	Rank By Percent Rating It Important	Percentage of Purchasers Checking		Dissatisfied: Somewhat Quite	Percent of Purchasers Dissatis- fied	Rank By Percent Dissatis- fied	
				Satisfied: Quite Somewhat	Dissatisfied: Somewhat Quite				
Television Sets, etc.	45.8	68.6	1	69.8	18.9	1.9	9.4	11.3	9
Radios, HiFi, Stereos	56.8	64.1	3	54.4	27.9	8.8	8.8	17.6	3
Home Movie Equipment	9.3	45.2	8	54.5	36.4	9.1	0	9.1	13
Pianos, Other Instruments	19.5	43.6	9	64.0	32.0	4.0	0	4.0	18
Records, Tape Recordings	62.7	52.8	5	51.4	37.8	8.1	2.7	10.8	10
Typewriters, calculators	57.6	50.0	6	55.2	31.3	6.0	7.5	13.5	8
Encyclopedias, Books, etc.	56.0	61.6	4	57.8	28.1	6.3	7.8	14.1	7
Arts and Crafts Equipment	39.0	43.3	10	58.7	37.0	4.3	0	4.3	17
Photographic Equipment	31.6	29.7	14	51.4	37.8	8.1	2.7	10.8	10
Home Workshop Equipment	33.3	46.2	7	69.2	20.5	10.3	0	10.3	12
Adult Sports Equipment	40.2	36.1	12	54.2	37.5	4.2	4.2	8.4	15
Boats, Motors, etc.	18.8	27.1	17	54.5	36.4	4.5	4.5	9.0	14
Rifles, Hand Guns, etc.	7.7	33.8	13	44.4	55.6	0	0	0	19
Camping Equipment	22.2	27.0	18	55.6	37.0	7.4	0	7.4	16
Luggage, Clothes Bags	44.4	28.8	15	51.9	30.8	5.8	11.5	17.3	4
Pool Tables, etc.	6.0	15.0	19	12.5	62.5	12.5	12.5	25.0	2
Children's Sports Equip- ment	23.1	36.8	11	50.0	35.7	7.1	7.1	14.2	6
Children's Games, Toys	34.2	27.4	16	48.8	36.6	7.3	7.3	14.6	5
Items not included above	4.1	65.8	2	33.3	0	0	66.7	66.7	1

*N = 119

TABLE 4

SATISFACTION/DISSATISFACTION RATINGS: CARS AND OTHER TRANSPORTATION DURABLES

Cars and Other Transportation Durable Categories	Use/Importance			Distribution of Responses				% Dissatisfied	
	Percent Having Made Purchase*	Percent Rating It Important	Rank By Percent Rating It Important	Percentage of Purchasers Checking		Dissatisfied:		Percent of Purchasers Dissatis-	Rank By Percent Dissatis-
				Satisfied: Quite Somewhat	Satisfied: Somewhat	Dissatisfied: Somewhat	Dissatisfied: Quite	fied	fied
New Car Purchase	30.2	97.0	5	40.5	27.0	18.9	13.5	32.4	4
New Truck, Panel Truck	4.3	100.0	1	85.7	14.3	0	0	0	10
Used Car Purchase	36.2	92.8	6	20.5	38.6	25.0	15.9	40.9	2
Used Pickup, Used Panel	3.5	100.0	1	33.3	33.3	16.7	16.7	33.4	3
Car or Truck Lease	16.4	36.6	11	42.9	33.3	19.0	4.8	23.8	5
Bicycle Purchase	41.4	54.1	10	44.9	34.7	6.1	14.3	20.4	6
Motorcycle Purchase	10.3	67.0	9	46.2	46.2	7.7	0	7.7	9
Travel Home Purchase	0.9	100.0	1	100.0	0	0	0	0	10
Airplane Purchase	0	0	12	--	--	--	--	--	--
Tires, Batteries, etc.	79.7	77.7	7	56.5	29.3	9.8	4.3	14.1	7
Parts for Home Auto Repair	43.4	77.4	8	53.1	38.8	8.2	0	8.2	8
Items not included above	2.7	100.0	1	33.3	0	33.3	33.3	66.6	1

*N = 119

TABLE 5

KENDALL'S TAU (C) CORRELATION COEFFICIENTS
DEMOGRAPHICS WITH SATISFACTION SCORES
HOUSING AND HOME FURNISHINGS

	Sex	Marital Status	Age	Employed	Number of Residents	Own/Rent Home	Education	Income
Single Family House	NS	-.2191	NS	-.2698	.2472	-.3743	NS	.2274
Condo. or Coop Unit	NS	.1305	-.1040	NS	NS	NS	NS	NS
Mobile Home	NS	NS	NS	NS	NS	NS	NS	NS
Vacation Home	NS	NS	NS	NS	NS	.1088	NS	.1924
Heating Installation	NS	-.1884	NS	-.1293	NS	-.1816	NS	.1198
Additions to Home	-.1671	-.1710	NS	NS	NS	-.2677	NS	.1626
Swimming Pool	-.2132	NS	NS	NS	NS	NS	NS	NS
"Do It Yourself" Projects	NS	-.1373	-.1644	-.1505	NS	NS	.1474	NS
Purchase of Land	NS	-.2063	NS	.2480	NS	NS	.1617	.2479
Draperies, Linens, etc.	NS	NS	NS	NS	.2183	NS	NS	NS
Electric Blankets, etc.	NS	NS	NS	-.1067	NS	NS	NS	NS
Baby Furniture	NS	-.2144	NS	NS	.1395	-.1730	.1526	.1060
Sofas, Chairs	NS	-.1434	NS	-.1681	.2107	NS	NS	NS
Din. Room, Kitchen Furniture	NS	NS	-.1045	-.1780	NS	NS	NS	NS
Bedroom Furniture	NS	NS	NS	-.1488	.1301	NS	.1549	NS
Outside Furniture	NS	NS	NS	NS	.1700	NS	NS	.1883
Floor Coverings	NS	-.1495	NS	-.2332	NS	-.1723	NS	.1061
Housewares, Tableware	NS	NS	-.2191	-.1044	NS	.1313	NS	NS
Lamps, Clocks, etc.	-.1512	NS	-.3300	-.1654	NS	.1516	NS	NS

Coefficients shown are significant at .05 level

TABLE 6

KENDALL'S TAU (C) CORRELATION COEFFICIENTS
 DEMOGRAPHICS WITH SATISFACTION SCORES
 HOME APPLIANCE AND PERSONAL CARE PRODUCTS

	Sex	Marital Status	Age	Employed	Number of Residents	Own/Rent Home	Education	Income
Refrigerators, Freezers	-.1031	-.2306	NS	-.2762	.2391	-.2955	NS	.2045
Air Conditioners, etc.	-.1275	-.1650	NS	-.1436	.2098	-.1620	NS	.1220
Ranges, Ovens, Grills	-.1928	NS	NS	-.3000	NS	NS	NS	.1495
Clothes Washers, Dryers	NS	-.1190	NS	-.1366	.1262	-.2208	NS	NS
Vacuum Cleaners, etc.	NS	NS	NS	-.2035	NS	NS	NS	NS
Sewing Machines, etc.	.1109	NS	NS	-.2608	NS	NS	NS	NS
Garbage Disposers, etc.	NS	NS	NS	NS	NS	-.1679	NS	NS
Small Kitchen Appliances	NS	NS	NS	NS	NS	NS	NS	NS
Electric Razors, etc.	-.2590	NS	-.1154	NS	NS	NS	NS	NS
Elec. Hairdryers, Curlers	NS	NS	-.2441	NS	.1620	NS	.1143	.1216
Exercise Machines, etc.	NS	NS	NS	NS	NS	NS	-.1459	NS
Vibrators, Massagers	NS	NS	NS	-.1642	NS	NS	NS	NS
Eyeglasses, Contact Lenses	NS	-.1026	NS	.1133	NS	NS	NS	.1525
Hearing Aids	-.1038	-.1294	NS	NS	NS	NS	NS	.1293
Medical Appliances, Devices	NS	-.1330	NS	-.1265	.1900	-.1488	NS	.1311

Coefficients shown are significant at .05 level

TABLE 7

KENDALL'S TAU (C) CORRELATION COEFFICIENTS
 DEMOGRAPHICS WITH SATISFACTION SCORES
 DURABLE ITEMS FOR ENTERTAINMENT, RECREATION, AND EDUCATION

	Sex	Marital Status	Age	Employed	Number of Residents	Own/Rent Home	Education	Income
Television Sets	NS	NS	NS	NS	NS	NS	NS	NS
Radios, HiFi, Stereos	NS	.1220	-.2905	NS	.1787	.1651	NS	NS
Home Movie Equipment	-.1206	NS	NS	NS	NS	NS	NS	NS
Pianos, Other Instruments	NS	NS	-.1625	-.1807	.1593	NS	NS	NS
Records, Tape Recordings	NS	.1123	-.3475	NS	.1822	.1250	NS	NS
Typewriters, Calculators	-.2474	-.1438	-.2866	-.1867	.1451	NS	.2521	.1310
Encyclopedias, Books, etc.	NS	-.1453	-.2410	NS	.1313	NS	.1382	NS
Arts and Crafts Equipment	NS	NS	-.1394	-.1868	NS	NS	NS	NS
Photographic Equipment	NS	NS	-.2003	-.1525	NS	NS	NS	NS
Home Workshop Equipment	-.2593	-.3330	NS	-.1956	.2257	-.1765	.2153	.2173
Adult Sports Equipment	-.2109	NS	-.3123	NS	.1575	.1555	NS	NS
Boats, Motors, etc.	-.1876	-.1368	NS	NS	NS	NS	NS	NS
Rifles, Hand Guns, etc.	NS	NS	-.1145	NS	NS	NS	NS	NS
Camping Equipment	-.1449	NS	NS	-.1917	.1100	NS	NS	NS
Luggage, Clothes Bags	NS	NS	-.2858	NS	NS	NS	NS	NS
Pool Tables, etc.	NS	-.1315	NS	-.1561	.1188	NS	NS	NS
Children's Sports Equipment	NS	-.1226	NS	-.1915	.3114	-.1453	NS	NS
Children's Games, Toys, etc.	NS	NS	NS	-.2455	.2185	-.1551	-.1359	NS

Coefficients shown are significant at .05 level

TABLE 8

KENDALL'S TAU (C) CORRELATION COEFFICIENTS
 DEMOGRAPHICS WITH SATISFACTION SCORES
 CARS AND OTHER TRANSPORTATION DURABLES

	Sex	Marital Status	Age	Employed	Number of Residents	Own/Rent Home	Education	Income
New Car Purchase	-.1799	-.3467	NS	-.1550	.2300	-.1613	.1403	.3665
New Truck, Panel Truck	-.1108	-.2010	NS	NS	NS	NS	NS	.1217
Used Car Purchase	-.1533	.1500	-.2269	-.1842	NS	.1331	NS	-.1108
Used Pickup, Used Panel	NS	-.1210	NS	NS	NS	NS	NS	NS
Car or Truck Lease	-.1181	-.1398	NS	NS	NS	NS	.1138	.1496
Bicycle Purchase	-.1229	-.1520	-.1554	NS	.3115	-.1884	.1575	.1070
Motorcycle Purchase	-.1297	NS	-.1507	NS	NS	NS	NS	NS
Travel Home Purchase	NS	-.1298	NS	NS	NS	NS	NS	NS
Airplane Purchase	NS	-.1055	NS	NS	NS	NS	NS	NS
Tires, Batteries, etc.	-.1929	NS	-.1276	NS	.2137	NS	NS	NS
Parts for Home Auto Repair	-.1305	.1054	-.2558	NS	.1075	NS	NS	-.1338

Coefficients shown are significant at .05 level

CONSUMER RESPONSE TO DISSATISFACTION WITH SERVICES AND INTANGIBLES

Ralph L. Day, Indiana University
Muzaffer Bodur, George Mason University

Abstract

Although most previous research on consumer dissatisfaction has emphasized complaints and redress-seeking, dissatisfied consumers may choose a number of other options including boycotting the brand, warning friends, or doing nothing at all. This paper describes a field study of consumer dissatisfaction and reports on patterns of dissatisfaction over 73 categories of services and intangible products. Summaries of reasons for dissatisfaction and post-dissatisfaction responses are also given.

Introduction

Increases in voiced complaints and redress-seeking have been both causes and effects of the dramatic growth of consumerism over the past decade. During most of this period a major source of information for consumer advocates, consumer protection agencies, and business firms has been tabulations of volunteered complaints. The users of complaint data have gradually become aware that the writers of complaint letters are not representative of all consumers and the particular items covered in complaint letters are not representative of the full scope of consumer concerns (Stokes, 1974). During the past few years, there has been a rapid increase in efforts to provide better quality data on consumer satisfaction and dissatisfaction through survey research (Lingoes and Pfaff, 1972; Pfaff, 1972; Warland, Herrmann, and Willits, 1975; Handy and Pfaff, 1975; Day and Landon, 1975) and efforts to develop better conceptualizations of satisfaction and dissatisfaction (Hunt, 1977; Day, 1977; Day and Landon, 1977). This paper seeks to make a contribution to knowledge and understanding of this increasingly important aspect of consumer behavior and to take a step toward the development of better methods of obtaining data on the post-dissatisfaction behavior of consumers.

The results reported here were obtained as part of a survey research project which utilized a set of questionnaires developed in 1976 under a research contract with the U.S. Office of Consumer Affairs. The basic objectives and overall structure of the study were similar to those of a previous project which was begun by the senior author at the Federal Trade Commission in 1974; however, many changes in the details of execution have resulted from field experience with previous versions of the questionnaire.

The unique feature of the study design is its comprehensive scope. A single study can obtain general indications of levels of satisfaction and dissatisfaction for almost 200 different product categories spanning the full range of consumer products and services. As respondents work through the sections of related product or service categories, they are asked periodically if they have had unsatisfactory experiences with items from any of the preceding set of categories. If the answer is yes, the respondent is asked to indicate the particular product or service with which they have had the most unsatisfactory experience, answer a battery of questions which probe for the sources of dissatisfaction, and provide information on any actions which might have been taken as a result of the unsatisfactory experience.

The survey was conducted in a single midwestern city in the fall of 1976. A two-stage probability sampling plan was used to select approximately 600 dwelling units. Self-

administered questionnaires were dropped off and picked up, and the overall response rate was approximately 80%. Four different questionnaires were assigned systematically over dwelling units. Three of these questionnaires covered different major categories of products or services; durable products, nondurable products, and services. The fourth questionnaire covered the entire range of products and services. The questionnaires which cover only one of the three major classes further divide the items into four sections and contain a probe for unsatisfactory experiences at the end of each of these. The combined questionnaire contains probes only at the end of each of the three major sections. Because of a concern that the longer combined questionnaire might have a lower completion rate, it was oversampled relative to the others. As a mild surprise, it turned out to have the highest completion rate. Usable responses were obtained from 125 respondents who got the separate services and intangibles questionnaires and from 171 respondents who were given the combined questionnaire.

Space does not permit the presentation and analysis of both sets of results or a complete comparative analysis of the effects of the increased opportunity to report unsatisfactory experiences on the pattern of responses. However, Table 1 reveals some of the differences in the pattern of reported dissatisfaction for the two questionnaires. A higher percentage of the respondents who had the separate services questionnaire named one or more instances of dissatisfaction (58.5% as compared to 39.8% of the respondents who had the combined questionnaire). Also, there were reports of dissatisfaction with items from a much higher percentage of the categories of services (67.1%) than the respondents who had the combined questionnaire (34.2%). Table 2 shows that the reported instances of dissatisfaction were more evenly divided over the four sections of the separate questionnaire, as compared with the corresponding groupings of the items in the combined questionnaire. These differences tend to support the expectation that separate questionnaires will provide a more sensitive indication of the pattern of consumer dissatisfaction.

Various alternative courses of action are available to consumers who have experienced dissatisfaction, ranging from doing nothing to suing a seller or manufacturer for millions of dollars in damages. The conceptual framework utilized in this study has been described elsewhere (Day and Landon, 1977). The dissatisfied consumer's options can be classified as follows:

- A. Take no action at all--forget the experience.
- B. Take some form of private action:
 - 1) change brands or supplier
 - 2) stop using the product class
 - 3) warn family or friends
- C. Take some form of public action:
 - 1) seek redress directly from the seller or manufacturer
 - 2) take legal action against the seller or manufacturer
 - 3) register a complaint with; the seller or manufacturer, a public consumer protection agency, or a private consumer organization.

These alternatives, except for the "do nothing" response, can be utilized by consumers in various combinations. Further discussion of this framework and some results

of previous studies are available elsewhere (Day and Landon, 1975).

Results and Analysis

This section will present and discuss the major findings from the separate services questionnaire. A brief description of the questionnaire will be helpful at this point. The 73 categories of services and intangible products are organized into four sections covering: repairs and general services; professional and personal services; financial services and insurance; and rentals, public transportation, and utilities. The sections were developed by grouping together those service items which are purchased and consumed in similar ways to form four sections of approximately equal length. The sections provide convenient units for retrospective probes for unsatisfactory experiences and the consumer's responses to them.

As in the other three questionnaires used in the present research project, the respondent works through each section of the services questionnaire one category at a time, first indicating whether or not any items from the category have been purchased within a specified time (the past two years in the case of services). Respondents who have made a purchase check a two point importance scale and a four point satisfaction scale ("always satisfied" to "always dissatisfied"). After they have completed the scales for all categories in a section, respondents are asked to look back over the section and indicate if they have had one or more experiences in which they were highly dissatisfied. If the answer is yes, the respondent then reports the number of unsatisfactory experiences with items from that section during the past two years and enters the number of the specific item which was "the most unsatisfactory." The respondent then checks any of 14 to 16 possible reasons for dissatisfaction with services which have contributed to dissatisfaction with that particular item. The reasons include statements such as "the service was provided in a careless and unprofessional manner." This is followed by a list of possible actions (including no action) that a dissatisfied consumer might take, based on the classifications of responses to dissatisfaction as outlined in the previous section. Data on the frequency of use of the various items and the frequency of "always dissatisfied" response from the satisfaction scale will be presented below, along with summaries and discussion of responses to the items found to be "the most unsatisfactory." Discussion of the full satisfaction/dissatisfaction scale and the importance ratings for both the separate services questionnaire and the services section of the combined questionnaire has been presented elsewhere (Day and Bodur, 1977).

Frequency of Dissatisfaction

One of the advantages of using survey research to study the consumer's response to dissatisfaction versus relying on volunteered complaints, is that it provides data on all users of a product or service and not just the complainers. In particular, data from a sample of the entire consumer population provide a frame of reference for putting the complaints of dissatisfied consumers in perspective. At the simplest level, an estimate of the total number of consumers allows the expression of the number of dissatisfied consumers as a percentage of the user population and permits the comparison of levels of dissatisfaction among users over different products or services. If one simply used the number of complaints as an indicator of dissatisfaction, then a service used by a small segment of the population would never be recognized as a consumer problem even if practically all of its users were dissatisfied. Conversely, a very widely used service might appear to be a major consumer problem

even though the rate of dissatisfaction with it is well below the average for all services.

The number of respondents who reported purchasing each of the 73 services included in the survey are shown in the first column of Tables 3-6, which cover the four sections of services. It can be seen that there is a wide range in the fractions of the sample using the various services. The number of instances of reported dissatisfaction for each kind of service in the four sections is shown in the second column of Tables 3-6. The percentage of users who were dissatisfied also varies widely. The highest rates of reported dissatisfaction for any of the 73 items occurred with services used by relatively small fractions of the sample; 21.1% of the users of employment agencies were dissatisfied and 20.0% of the users of nursing homes reported dissatisfaction (Table 4). These services were used by 15.2% and 4.0% of the sample, respectively. Other services with high rates of dissatisfaction among users but relatively small numbers of users included furniture rental services (Table 6) and septic tank services (Table 3). The highest percentage of dissatisfied respondents in any widely used category was 18.8% for auto repairs with 80.8% of the sample reporting a purchase within the past two years (Table 2). Auto repairs has headed the list of volunteered complaints for years, but some items with substantial rates of dissatisfaction but smaller user bases have not ranked high on complaint lists.

To determine if the extreme negative responses on the satisfaction scale (always dissatisfied) are predictive of the frequency of reported dissatisfaction provided at the end of each section, the two sets of responses were compared. Column 3 in Tables 3-6 contains the number of "always dissatisfied" responses for each item in the section. Spearman rank correlation coefficients between the two sets of responses in each section were as follows: repairs and general services (Table 3), $r_s = .91$; professional and personal services (Table 4), $r_s = .83$; financial services and insurance (Table 5), $r_s = .93$; rentals, public transportation, and utilities (Table 6), $r_s = .96$. All correlations were significant beyond the .001 level, suggesting that the responses reflecting extreme dissatisfaction on satisfaction/dissatisfaction scales for particular categories are predictive of reports of dissatisfaction.

Reasons for Dissatisfaction

Respondents who indicated that they had been highly dissatisfied with a particular service considered a list of possible reasons for being dissatisfied and were instructed to check all of the reasons which they felt applied to their unsatisfactory experience. If they checked multiple reasons, respondents were asked to indicate which one of the items was the most important. The reasons varied slightly in number and content over the four sections of services to reflect differences in the nature of the services and the circumstances of purchase and use. The wording of the statements, the number of times each was checked, and the number of times each was named as most important are given for each of the four sections in Tables 7-10. The number of instances of reported dissatisfaction varies over the sections with "repairs and general services" showing considerably more causes of dissatisfaction than the other sections. The average number of reasons given in a particular instance of dissatisfaction was very consistent over the four sections, and the overall average number of reasons given was 2.62.

The great majority of the reasons given for dissatisfaction were directly related to the quality of the supplier's performance. Because of the differences in the nature of the particular service items, several different statements were related to various aspects of the performance

of the service. By far the most frequently mentioned reason for dissatisfaction was "The service was rendered in a careless, unprofessional manner." The percentage of those respondents reporting an instance of dissatisfaction who checked this item ranged from 58.3% for rentals; public transportation, and utilities to 84.6% for professional and personal services. It was also checked as the most important of the reasons for all four sections. The stronger indication of gross incompetence with harmful results was the second most frequently checked reason for dissatisfaction in the repairs and general services and the professional and personal services sections. Some reasons related to quality which were relevant for only one or two of the sections were important in those sections. For example, in the repairs and general services section, the third and fourth most frequent responses were "performance of the item was worse after the repairs than before" and "the quality of parts or materials was inferior."

The various possible reasons for dissatisfaction relating to marketing practices as opposed to "service quality" were checked very infrequently. The statements relating to misleading advertising, deceptive practices of salesmen, credit practices, and warranties were also checked infrequently. While there was some degree of concern about prices or fees, these were completely overshadowed by concerns with the quality of performance. A fraudulent practice related to the price or fee, "I was charged for work that was not done," was indicated in 19.2% of the instances in the "professional and personal services" section and in 19.4% of the instances in the "rentals, public transportation, and utilities" section. Similar reasons in the other two sections were checked somewhat less frequently.

The results shown in Tables 7-10 strongly suggest that the respondents who reported dissatisfaction with services were much more strongly concerned with performance or "service quality" issues than with marketing practices. Previous studies of tangible products have tended to show that aspects other than product quality play a more important role in voiced complaints. Although some of the data from the durable products and nondurable products sections of the present study have not been analyzed, the results of a previous study using an earlier version of the "combined" instrument used in this study suggested that marketing practices and price-related issues are somewhat more frequently mentioned for both durables and nondurables than for services (Day and Landon, 1976). This may be a reflection of the more direct and personal relationships involved in the purchase and use of many services than in the purchase and use of tangible products. Hopefully, the completion of the analysis of the present project and additional future research on the marketing of services will shed light on this and many other issues relating to services.

Post-Dissatisfaction Responses

In 29 of the 133 reported instances of dissatisfaction (21.8%), the respondent indicated that no action was taken. Although a detailed probe for reasons for these respondents' nonbehavioral response to dissatisfaction was not feasible, respondents were asked to indicate which of four possible reasons best explained why they had taken no action. These responses are shown in Table 11. None of the respondents checked the "procrastinator's response" of "I wanted to do something about it but never got around to it." The modal response was "I didn't think anything I could do would make any difference." This appears to reflect either a defeatist attitude or a pessimistic appraisal of the chances for successfully obtaining redress. With the possible exception of the "not worth it" response, it seems likely that better information and encouragement to seek redress might lead to a reduction in the number of no-action responses.

In 104 of the 133 instances, the respondent reported taking some action. A summary of reported actions is given in Table 12. A total of 250 actions were reported, an average of 2.4 actions per instance of reported dissatisfaction. These are classified as private actions, redress seeking, and complaining. The total number of actions of each of the three types are shown in the right hand column in Table 12. The distribution of the specific responses within each of the three types is shown in the body of Table 12. Private actions accounted for 46.4% of total actions and 81% of the private actions involved either changing to another supplier of the service or warning others. The remaining 19% reported that they had decided to quit using that service altogether. The public actions were equally split between redress seeking and complaining. Redress seekers primarily sought "replacement" (redoing) of the service or a financial adjustment. An overwhelming majority of the public complaints (80.6%) were directed to the company or professional person providing the service.

The results presented in Tables 10 and 11 suggest that while dissatisfied consumers take no action in connection with a substantial number of cases, approximately 80% of such instances do result in some form of action. More than 45% of the reported actions were of a "private" nature of the type that would never be brought to the attention of business firms, consumer advocates, or consumer protection agencies.

Summary and Conclusions

This paper has presented some of the results from a broader study of consumer satisfaction/dissatisfaction and post-dissatisfaction response. The focus was on the post-evaluation responses from a comprehensive study of consumer services, using the data from a probability sample of 125 households in a single midwestern city. The data include the total number of users of each of 73 categories of services as well as the number of instances of reported dissatisfaction for each category. This allows the evaluation of relative levels of expressed dissatisfaction among the various types of services. A classification of responses into nonbehavioral (no action), private actions, redress-seeking, and public complaining was applied to the results. It was found that in more than 20% of the instances of expressed dissatisfaction, no action was taken and that 46.5% of reported actions were of a private nature, such as boycotting the supplier or warning friends.

These results suggest that more complete and more readily interpretable data about levels of consumer dissatisfaction can be obtained by survey research than by keeping records of public complaints. The results also suggest that comprehensive studies of the entire range of consumer products and services are feasible and can provide useful data of a type that is not readily obtainable in any other way. The comprehensive survey shows promise as a diagnostic tool for identifying areas in which more detailed data are needed. It also offers considerable promise as a tool for consumer behavior theorists who wish to develop more complete models of the consumer's purchase, consumption, and evaluation process. Although the results reported here were from a rather small sample from a single city and cannot be freely generalized, they suggest the need for a large national study as a starting point for the development of a comprehensive data base on consumer satisfactions, dissatisfactions, and post-dissatisfaction responses.

References

Ralph L. Day, "Extending the Concept of Consumer Satisfaction," in William D. Perrault, ed., Advances in Consumer Research, Volume IV, Atlanta: Association for Consumer Research, 1977, 149-154.

Ralph L. Day and Muzaffer Bodur, "A Comprehensive Study of Satisfaction with Consumer Services and Intangibles," in Ralph L. Day, ed., Consumer Satisfaction, Dissatisfaction, and Complaining Behavior, Division of Business Research, Indiana University, 1977, 64-74.

Ralph L. Day and E. Laird Landon, "Survey Data on Consumer Complaints for Consumer Protection Policy Makers," Proceedings, Midwest AIDS Conference, 1975, 40-44.

Ralph L. Day, "Collecting Comprehensive Consumer Complaint Data by Survey Research," in Beverlee B. Anderson, ed., Advances in Consumer Research, Volume III, Chicago: Association for Consumer Research, 1976, 263-268.

Ralph L. Day, "Toward a Theory of Consumer Complaining Behavior," in Arch Woodside, Jagdish Sheth and Peter Bennett, eds., Foundations of Consumer and Industrial Buying Behavior, American Elsevier, 1977, 425-437.

Charles R. Handy and Martin Pfaff, "Consumer Satisfaction with Food Products and Marketing Services," Economic Research Service, U.S. Department of Agriculture, Agricultural Economic Report No. 281, Washington, D.C., March 1975.

H. Keith Hunt, editor, Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction, Marketing Science Institute, 1977.

James C. Lingoies and Martin Pfaff, "The Index of Consumer Satisfaction: Methodology," Proceedings, Association for Consumer Research, 1972, 689-712.

Anita B. Pfaff, "An Index of Consumer Satisfaction," Proceedings, Association for Consumer Research, 1972, 713-737.

Raymond C. Stokes, "Consumer Complaints and Consumer Dissatisfaction," Speech at Food and Drug Law Institute, Phoenix, Arizona, April 1974.

Rex H. Warland, Robert O. Herrmann and Jane Willits, "Dissatisfied Consumers: Who Gets Upset and Who Takes What Action," Journal of Consumer Affairs, Winter 1975, 148-163.

TABLE 1

COMPARISON OF RESULTS WITH THE TWO QUESTIONNAIRES		
	Separate Services	Services Section In Combined
Sample Size	125	171
Number Reporting Dissatisfaction	73 (58.4%)	68 (39.8%)
Number Naming Only One Item	34 (27.2%)	68 (39.8%)
Number of Dissatisfied Taking no Action	12 (9.6%)	18 (12.5%)
Number of Service Categories Named	49 of 73 (67.1%)	25 of 73 (34.2%)

TABLE 2

INSTANCES OF DISSATISFACTION OVER SUBSECTIONS

Sections	Services Only	Combined
	Respondents = 73 Responses = 133	Respondents = 68 Responses = 68
Repairs and General Services	48 (36.1%)	22 (32.4%)
Professional and Personal Services	26 (19.5%)	6 (8.8%)
Financial Services and Insurance	23 (17.3%)	7 (10.3%)
Rentals	36 (28.8%)	33 (48.5%)
TOTAL	133 (100%)	68 (100%)

$\chi^2 = 9.47$
Significant at .025 level

TABLE 3

FREQUENCY OF USE AND DISSATISFACTION: REPAIRS AND GENERAL SERVICES

Categories of Repairs and General Services	Users of this Service Category		Users Reporting Dissatisfaction		Users "Always Dissatisfied"*	
	No	%	No	%	No	%
Auto repairs and service	101	80.8	19	18.8	9	8.9
Photographic services and processing	93	74.4	0	0.0	1	1.1
Laundry and dry cleaning	92	73.6	1	1.1	3	3.3
Parcel delivery	84	67.2	0	0.0	4	4.8
Do-it-yourself laundry, dry cleaning	80	64.0	2	2.5	3	3.8
TV, radio, stereo repairs	76	60.8	4	5.3	2	2.6
Services of mail order firms	65	52.0	5	7.7	5	7.7
Watch, clock and jewelry repairs	60	48.0	0	0.0	2	3.3
Plumbing, carpentry, home repairs	53	42.4	3	5.7	2	3.8
Heating, air cond. repairs and service	48	38.4	2	4.2	1	2.1
Appliance repairs (except TV, radio, stereo)	43	34.4	4	9.3	3	7.0
Carpet cleaning, window washing, home care	37	29.6	0	0.0	1	2.7
Yardwork and lawn care services	27	21.6	1	3.7	2	7.4
Furniture reupholstery, refinishing	27	21.6	1	3.7	1	3.7
Moving and storage	25	20.0	1	4.0	3	12.0
Installation of siding, roofing, painting	25	20.0	2	8.0	1	4.0
Remodeling, decorating, landscaping	20	16.0	2	10.0	2	10.0
Domestic help, maid service	11	8.8	0	0.0	1	9.1
Cesspool, septic tank service	7	5.6	1	14.3	1	14.3
Water softening service	6	4.8	0	0.0	0	0.0

*Data in this column is from the item-by-item satisfaction scale and shows the number of users of the service who checked "always dissatisfied." The Spearman rank correlation between the instances of reported dissatisfaction and the number of "always dissatisfied" responses is .91, significant at the .001 level.

TABLE 4

FREQUENCY OF USE AND DISSATISFACTION: PROFESSIONAL AND PERSONAL SERVICES

Categories of Professional and Personal Services	Users of this		Users Reporting		Users "Always	
	Service Category		Dissatisfaction		Dissatisfied"*	
	No	%	No	%	No	%
Dentists and dental technicians	108	86.4	4	3.7	2	1.9
Medical doctors, nurses in office or home	107	85.6	5	4.7	4	3.7
Barber and beauty shops, health spas	95	76.0	3	3.2	3	3.2
Medical doctors, nurses in hospitals	85	68.0	5	5.9	3	3.5
Optometrists, ophthalmologists	83	66.4	1	1.2	1	1.2
Veterinarians, animal hospitals	64	51.2	0	0.0	1	1.6
Travel agencies	43	34.4	0	0.0	1	2.3
Lawyers	39	31.2	2	5.1	5	12.8
Income tax preparation services	38	30.4	1	2.6	2	5.3
Funeral homes, mortuaries, cemeteries	23	18.4	0	0.0	1	4.3
Psychologists, marriage counselors	19	15.2	0	0.0	1	5.2
Employment agencies	19	15.2	4	21.1	6	31.6
Osteopaths, chiropractors, physical therapists	15	12.0	0	0.0	0	0.0
Nursing homes and rest homes	5	4.0	1	20.0	1	20.0
Computer dating, matching	4	3.2	0	0.0	0	0.0
Architects, home designers	2	1.6	0	0.0	0	0.0
Home security agencies, private detectives	0	0.0	0	0.0	0	0.0

*Data in this column is from the item-by-item satisfaction scale and shows the number of users of the service who checked "always dissatisfied." The Spearman rank correlation between the instances of reported dissatisfaction and the number of "always dissatisfied" responses is .83, significant at the .001 level.

TABLE 5

FREQUENCY OF USE AND DISSATISFACTION: FINANCIAL SERVICES AND INSURANCE

Categories of Financial Services and Insurance	Users of this		Users Reporting		Users "Always	
	Service Category		Dissatisfaction		Dissatisfied"*	
	No	%	No	%	No	%
Banks and trust companies	114	91.2	7	6.1	3	2.6
Hospitalization, surgical-medical insurance	84	67.2	1	1.2	0	0.0
Savings and loan associations	78	62.4	0	0.0	0	0.0
Auto collision insurance	75	60.0	3	4.0	1	1.3
Credit card services	75	60.0	3	4.0	1	1.3
Liability insurance	63	50.4	1	1.6	1	1.6
Life insurance	63	50.4	2	3.2	1	1.6
Health, income protection insurance	56	44.8	0	0.0	0	0.0
Homeowners or renters insurance	55	44.0	1	1.8	2	4.2
Group insurance plans	48	38.4	2	4.2	0	0.0
Fire, theft, property damage insurance	48	38.4	1	2.1	1	0.0
Credit unions	43	34.4	1	2.3	1	2.3
Stockbrokers, investment counselors, security dealers	19	15.2	0	0.0	1	5.3
Small loan companies	17	13.6	1	5.9	1	5.9

*Data in this column is from the item-by-item satisfaction scale and shows the number of users of the service who checked "always dissatisfied." The Spearman rank correlation between the instances of reported dissatisfaction and the number of "always dissatisfied" responses is .93, significant at the .001 level.

TABLE 6

FREQUENCY OF USE AND DISSATISFACTION: RENTALS, PUBLIC TRANSPORTATION AND UTILITIES

Categories of Rentals, Public Transportation and Utilities	Users of This Service Category		Users Reporting Dissatisfaction		Users "Always Dissatisfied"*	
	No	%	No	%	No	%
U.S. Postal Service	122	97.6	6	4.9	2	1.6
Local telephone company	119	95.2	6	5.0	5	4.2
Local electric company	112	89.6	3	2.7	3	2.7
Local garbage and trash collection	95	76.0	1	1.1	1	1.1
Local water company	89	71.2	1	1.1	1	1.1
Local natural gas company	86	68.8	2	2.3	1	1.2
Rooms in hotels, boarding houses	79	63.2	1	1.3	2	2.5
Major schedule airline service	71	56.8	2	2.8	1	1.4
Local cable television company	68	55.4	3	4.4	4	5.9
Apartment rental	64	51.2	7	10.9	8	12.5
Local taxi service	45	36.0	3	6.7	1	2.2
Local bus service	45	36.0	0	0.0	0	0.0
Car, truck, trailer, airplane, and boat rentals	38	30.4	1	2.6	2	5.3
House rental	33	26.4	6	18.2	2	6.1
Inter-city bus service	31	24.8	1	3.2	0	0.0
Air commuter, air charter service	29	23.2	1	3.4	0	0.0
Equipment, tools, party supplies rentals	29	23.2	0	0.0	0	0.0
Local L.P. gas, fuel oil suppliers	26	20.8	1	3.8	0	0.0
Passenger service on trains	20	16.0	3	15.0	2	10.0
Mobile home rental	6	4.8	2	33.3	0	0.0
Furniture, appliance rentals	6	4.8	0	0.0	1	16.7
Uniform rental, diaper and linen services	5	4.0	0	0.0	0	0.0

*Data in this column is from the item-by-item satisfaction scale and shows the number of users of the service who checked "always dissatisfied." The Spearman rank correlation between the instances of reported dissatisfaction and the number of "always dissatisfied" responses is .96, significant at the .001 level.

TABLE 7

REASONS FOR DISSATISFACTION: REPAIRS AND GENERAL SERVICES

Reasons for Being Dissatisfied	Total Number of Mentions		Named Most Important	
The service was provided in a careless, unprofessional manner;	31	64.6%	10	43.5%
Services were rendered in an incompetent manner with very harmful results;	18	37.5	2	8.7
Performance of the item was worse after the repairs than before;	17	35.4	2	8.7
The quality of parts or materials was inferior;	15	31.3	1	4.3
The service was not completed in the agreed time;	13	27.1	1	4.3
Results fell short of those claimed by ads;	9	18.8	1	4.3
I was charged for work that was not done;	8	16.7	4	17.4
The fee was much higher than the amount agreed in advance;	6	12.5	0	0.0
I was charged for parts that were not furnished;	4	8.3	0	0.0
Unauthorized repairs were made and charged to me;	2	4.2	2	7.0
I was tricked by a salesman into buying services I did not want;	2	4.2	0	0.0
I was harassed by bill collectors;	1	2.1	0	0.0
The warranty did not cover everything that went wrong;	1	2.1	0	0.0
Credit terms were misrepresented to me;	0	0.0	0	0.0
	127		23	

TABLE 8

REASONS FOR DISSATISFACTION: PROFESSIONAL AND PERSONAL SERVICES				
Reasons for Being Dissatisfied	Total Number of Mentions		Named Most Important	
The service was provided in a careless, unprofessional manner;	22	84.6%	9	40.9%
Services were rendered in an incompetent manner with very harmful results;	10	38.5	4	18.2
Things were worse after the service than before;	8	30.8	2	9.1
The professional advice I paid for was incorrect and caused me substantial losses;	6	23.1	3	13.6
I was charged for work that was not done;	5	19.2	1	4.5
Results fell short of those claimed by ads;	4	15.4	1	4.5
The quality of materials, or medicines which were furnished was inferior;	4	15.4	0	0.0
The service was not completed at the agreed time;	2	7.7	2	9.1
I was charged for materials or medicines that were not furnished;	1	3.4	0	0.0
The fee was much higher than the amount agreed in advance;	1	3.4	0	0.0
A professional confidence was violated to my embarrassment or injury;	1	3.4	0	0.0
I was tricked by a salesman into buying the services I did not want;	1	3.4	0	0.0
I was harassed by bill collectors;	1	3.4	0	0.0
Credit terms were misrepresented to me;	1	3.4	0	0.0
The warranty did not cover everything that went wrong;	0	0.0	0	0.0
	<u>67</u>		<u>22</u>	

TABLE 9

REASONS FOR DISSATISFACTION: FINANCIAL SERVICES AND INSURANCE				
Reasons for Being Dissatisfied	Total Number of Mentions		Named Most Important	
The service was provided in a careless unprofessional manner;	16	69.6%	1	5.6%
Many mistakes were made in my account;	10	43.5	4	22.2
Services were rendered in an incompetent manner with very harmful results;	5	21.7	3	16.7
The service was not completed at the agreed time;	5	21.7	1	5.6
I was charged for services that were not performed;	4	17.4	1	5.6
I was harassed by bill collectors;	4	17.4	1	5.6
The benefits did not cover all of the expenses as claimed;	3	13.0	2	11.1
I was unfairly refused credit or other financial services;	3	13.0	1	5.6
Results fell short of those claimed by ads;	2	8.7	1	5.6
The company refused to pay a valid claim;	2	8.7	1	5.6
The fee was much higher than the amount agreed in advance;	2	8.7	0	0.0
A professional confidence was violated to my embarrassment or injury;	2	8.7	0	0.0
Credit terms were misrepresented to me;	1	4.3	1	5.6
My insurance policy was cancelled without justification;	1	4.3	1	5.6
The professional advice I paid for was incorrect and caused me substantial losses;	1	4.3	0	0.0
I was tricked by a salesman into buying services; insurance or other intangibles I did not want;	1	4.3	0	0.0
	<u>62</u>		<u>18</u>	

TABLE 10

REASONS FOR DISSATISFACTION: RENTALS, PUBLIC TRANSPORTATION AND UTILITIES

Reasons for Being Dissatisfied	Total Number		Named Most	
	of Mentions		Important	
The service was provided in a careless, unprofessional manner;	21	58.3%	7	25.9%
The quality of the service was inferior;	16	44.4	3	11.1
The service was not performed in a reasonable time;	15	41.7	7	25.9
The premises and/or items that were leased or rented to me were not in good condition;	12	33.3	4	14.8
The quality of the item I rented was inferior;	8	22.2	2	7.4
I was charged for work that was not done;	7	19.4	1	3.7
Things that belonged to me were damaged or lost;	4	11.1	0	0.0
Results fell short of those claimed by ads;	3	8.3	0	0.0
I was charged for materials or parts that were not furnished;	2	5.6	1	3.7
The fee was much higher than the amount agreed in advance;	2	5.6	1	3.7
I was harassed by bill collectors;	2	5.6	1	3.7
Unauthorized repairs were made and charged to me;	1	2.8	0	0.0
I was tricked by a salesman into buying services I did not want;	0	0.0	0	0.0
Credit terms were misrepresented to me;	0	0.0	0	0.0
The warranty did not cover everything that went wrong	0	0.0	0	0.0
	<u>93</u>		<u>27</u>	

TABLE 11

REASONS DISSATISFIED RESPONDENTS GAVE FOR TAKING NO ACTION

"The one particular reason which best explains why you decided not to do anything"	Replies	
	No.	%
I didn't think it was worth the time and effort	10	27.8
I wanted to do something but I didn't get around to it	0	0.0
I didn't think anything I could do would make any difference	18	50.0
I didn't know what I could do about it, or where I could get help	8	22.2
TOTAL	<u>36</u>	<u>100.0</u>

TABLE 12

SUMMARY OF ACTIONS TAKEN IN RESPONSE TO DISSATISFACTION*

Response	Type of Action			No. of Actions	
I. PRIVATE ACTION:		<u>Private Actions:</u>		116	46.4%
		<u>No.</u>	<u>%</u>		
		Decided not to buy the service again;	22	19.0	
		Quit the company or professional person;	44	37.9	
		Warned family and friends;	50	43.1	
	TOTAL	116	100.0		
II. PUBLIC ACTION: Sought Redress		<u>Redress Actions:</u>		67	26.8%
		<u>No.</u>	<u>%</u>		
		Requested service be performed in correct way;	33	49.2	
		Asked for refund or adjustment of fee;	26	38.8	
		Contacted a lawyer or otherwise took legal action	7	10.5	
		Requested that the claim be paid;	1	1.5	
	TOTAL	67	100.0		
III. PUBLIC ACTION: Complained		<u>Complaint Actions:</u>		67	26.8%
		<u>No.</u>	<u>%</u>		
		To the company of professional person;	54	80.6	
		To a professional association;	3	4.5	
		To the Better Business Bureau;	1	1.5	
		To a government agency or official	5	7.5	
		To a consumer advocate or organization;	4	5.9	
		TOTAL	67	100.0	250

*In 29 of the 133 cases of reported dissatisfaction, no action was taken. This table refers to the 104 instances in which one or more actions were taken.

CATASTROPHE THEORY AS A MODEL
FOR DESCRIBING CONSUMER BEHAVIOR

Terence A. Oliva, Louisiana State University
Alvin C. Burns, Louisiana State University

Abstract

Consumer behavior models do not have mechanisms to handle divergent and discontinuous behavior. Rene Thom's Catastrophe Theory seems to be a promising means for modeling such phenomena. This paper describes the cusp catastrophe and illustrates how it can describe consumer adoption and consumer complaining behavior.

function which relates various combinations of control variables (independent variables) equal to zero (Starobin, 1976). Thom has shown that only seven models are needed to describe all possible abrupt transitions in situations that are controlled by up to four independent variables (Zeeman, 1976). The mathematical formulations of these models are contained in Table 1.

Introduction

It is an understatement to contend that the behavioral scientist encounters frustration when he attempts to model the phenomena under his study. Unlike his counterparts in the physical sciences, the behavioral scientist finds his subject matter highly resistant to straightforward conceptualization. Simple models cannot possibly cope with the myriads of multiple determinants, complex interactions, non-linear relationships, discontinuities and seemingly unsystematic variations.

Consumer behaviorists have confronted these problems for some years now. While we believe we have transcended simplistic "black box" models and can portray the behavior of a consumer in the context of ladder or stairs concepts, decision trees, and elaborate flow diagrams, the behavior inevitably escapes the boundaries of the model when we attempt to describe it in appreciable detail. One possible reason for the disappointing experiences thus far may be in the implicit assumption in most consumer behavior models that behavior follows a smooth transition through and between stages in the process. As a more analytical posture is assumed, however, it soon becomes evident that such smooth transitions are rare. More commonly observed are sudden shifts, stalls, and divergent behavior. A mechanism for describing these phenomena must be included if our models are to faithfully depict the process.

Catastrophe Theory

As indicated above, a major shortcoming of consumer behavior models has been the inability to handle abrupt (catastrophic) changes in behavior. Modeling of what might be called "misbehaving phenomena" has defied mathematical analysis for a long time. The reason for this is that such analysis usually requires some form of continuum. Relatively recently, however, French mathematician Rene Thom (1972) using a branch of mathematics known as Topology developed some ideas which generate models that can handle messy phenomena.

Topology is a kind of rubber sheet geometry that allows elastic movements (Arnold, 1962). It is the branch of mathematics that deals with the fundamental properties of literally anything; and, as such, focuses on invariance through transformation (Arnold, 1962; Barr, 1964). Thom uses that part of topology that deals with multi-dimensional surfaces because of its ability to represent underlying forces in nature under different states of equilibrium (Zeeman, 1976).

Structurally, catastrophe models are topological surfaces generated by setting the first derivative of a

TABLE 1
Rene Thom's Seven Elementary Catastrophes

CATASTROPHE	CONTROL DIMENSIONS	BEHAVIOR DIMENSIONS	FUNCTION	FIRST DERIVATIVE
FOLD	1	1	$\frac{1}{3}x^3 - ax$	$x^2 - a$
CUSP	2	1	$\frac{1}{4}x^4 - ax - \frac{1}{2}bx^2$	$x^3 - a - bx$
SWALLOWTAIL	3	1	$\frac{1}{5}x^5 - ax - \frac{1}{2}bx^2 - \frac{1}{3}cx^3$	$x^4 - a - bx - cx^2$
BUTTERFLY	4	1	$\frac{1}{6}x^6 - ax - \frac{1}{2}bx^2 - \frac{1}{3}cx^3 - \frac{1}{4}dx^4$	$x^5 - a - bx - cx^2 - dx^3$
HYPERBOLIC	3	2	$x^3 + y^3 + ax + by + cxy$	$3x^2 + a + cy$ $3y^2 + b + cx$
ELLIPTIC	3	2	$x^3 - xy^2 + ax + by + cx^2 + cy^2$	$3x^2 - y^2 + a + 2cx$ $-2xy + b + 2cy$
PARABOLIC	4	2	$x^2y + y^4 + ax + by + cx^2 + dy^2$	$2xy + a + 2cx$ $x^2 + 4y^3 + b + 2dy$

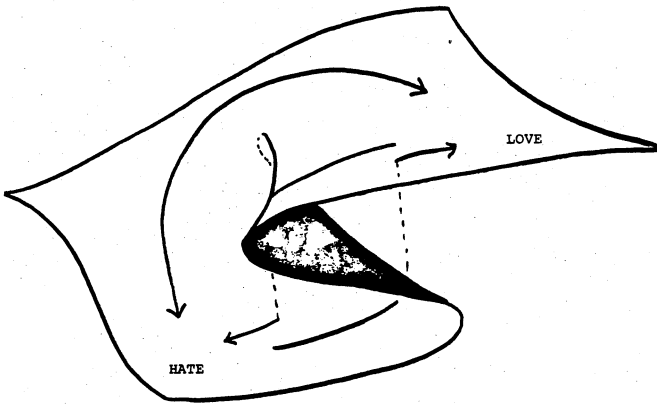
SEVEN ELEMENTARY CATASTROPHES describe all possible discontinuities in phenomena controlled by no more than four factors. Each of the catastrophes is associated with a potential function in which the control parameters are represented as coefficients (a, b, c, d) and the behavior of the system is determined by the variables (x, y). The behavior surface in each catastrophe model is the graph of all the points where the first derivative of this function is equal to zero or, when there are two first derivatives, where both are equal to zero.

(From Zeeman, 1976)

The way in which catastrophe models are able to describe phenomena that diverge, are subject to sudden shifts, and are discontinuous is through folds and pleats in the surfaces that relate different equilibrium states. Love and hate, for example, represent two different states.

Transition from one to the other does not necessarily progress along a unidimensional continuum. At the same time, it would be incorrect to say that they occupy two disconnected continua such that an individual cannot go from one to the other. A catastrophe model would relate the two by means of a surface fold. (See Figure 1)

FIGURE 1
A Catastrophe Behavior Surface



The movement from hate to love or vice versa can occur a number of ways on the surface. A smooth transition develops if the process is such that the individual misses the fold. On the other hand, an abrupt transition occurs when the individual "falls off" one surface to the other by crossing the fold. Thus the theory describes equally well both the persistent suitor who slowly wins his heartthrob's affection and the wife who finds her cheating husband in bed with another woman.

Some of the situations where catastrophe models have been used are: the development of hostilities between nations, cathartic release from self-pity, aggression in dogs, behavior of the stock market, buckling of an elastic beam, anorexia nervosa, territoriality among reef fish (Zeeman, 1976), population dynamics (Starobin, 1976), and collective bargaining (Oliva and Capdevielle, 1977). All the preceding situations have conditions where abrupt transitions occur, making the use of other modeling techniques virtually impossible on anything but a trivial level.

Perhaps the best way to understand what catastrophe theory is all about is by close examination of one of its models. The model we have chosen is the cusp catastrophe since it is easily understood and has been the most widely used of the models. A Cusp Model is generated by setting the first derivative of the following function equal to zero for all combinations of the variables a, b, and x, where a and b are two control dimensions, and x represents the system's behavior dimension:

$$f(a, b, x) = 1/4x^4 - ax - 1/2 bx^2 \quad (1)$$

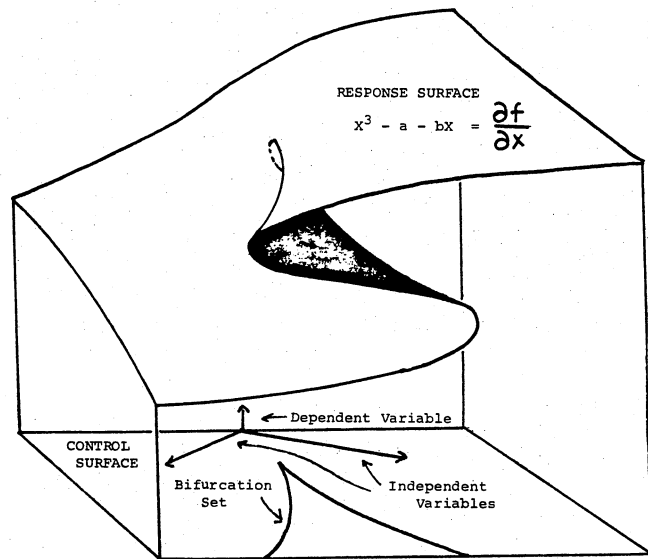
$$\frac{\partial f}{\partial x} = 0 = x^3 - a - bx \quad (2)$$

The response surface generated is peculiar (as indicated earlier) in that it has a fold in it. (See Figure 2) By projecting the fold onto the control surface, an area of overlap known as the bifurcation set is defined. It is domain where two modes of behavior are possible depending on the initial conditions.

For a situation to be suitable for description by the cusp catastrophe, five conditions are requisite:

bimodality, sudden transitions, hysteresis, unaccessibility, and divergence (Zeeman, 1976).

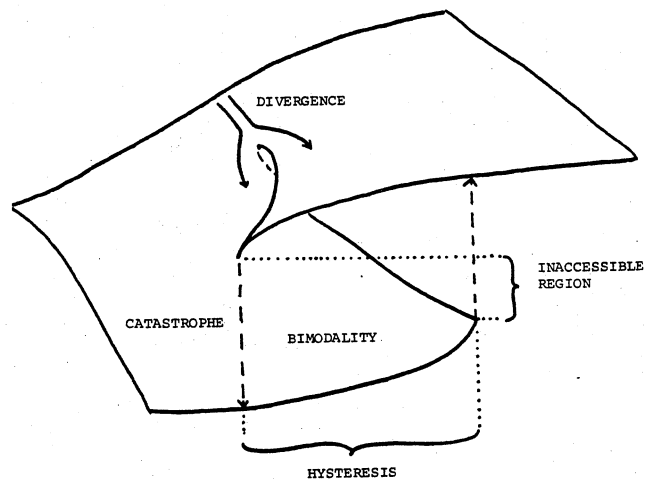
FIGURE 2
A Cusp Catastrophe



(From Zeeman, 1976)

Referring to Figure 3, we can see how the five conditions stated above are related by the model. It should be apparent from the diagram that the fold is the focal point of the model. Taking each condition in the abstract we can consider how the model operates in general.

FIGURE 3
A Cusp Catastrophe's 5 Requirements



(Zeeman, 1976)

Divergent behavior is accounted for since as one moves out from the edge toward the singularity, which is the starting point of separation between the two surfaces, the system is forced to adopt one of the two opposing behaviors (e.g., love or hate; attack or flee). The trajectory that will be followed will depend on the values of the control variables (state of the system) just prior to reaching the singularity.

Bimodality is created in the model by the top and bottom surfaces of the fold. It represents the domain where the system tends to hold equal amounts of both control variables at the same time. In a love/hate situation, the bifurcation set would represent the area of ambivalence.

The middle sheet of the fold represents the inaccessible area which is the least likely behavior. In the love/hate situation, this would be the condition of neutrality or non-emotion. To see why it is inaccessible, consider the following. If we hold control variable b constant at a value of 3 and increase the value of a , the system trajectory will proceed in a smooth fashion along the bottom surface until it reaches $(2,3,-1)$ where it will jump to $(2,3,2)$ on the upper surface and continue smoothly. Consequently, there is no means for the behavior to reach the middle sheet.

Sudden transitions (catastrophes) are accounted for as suggested in the preceding paragraph at the edges of the fold. That is, as the system moves along the surface toward the pleat at some point a small increase in the control variable will cause a sudden shift in behavior. To use Zeeman's (1976) example, if an angry dog is made more fearful, at some point his behavior will abruptly change from attack to retreat; conversely, if a fearful dog is progressively enraged, at some point he will stop retreating, turn and attack. It is the ability of the model to represent sudden shifts that gives it great advantage over other conceptualizations.

The fact that the catastrophes do not occur at the same place but depend on the prior state of the system incorporates lags into the model which match real-world phenomena. For example, if a person starts out "in love" and is made to increasingly dislike the other, the complete changeover to hate will not occur at the same point as a person who starts out hating another. That is, a person in love endures more grief from a second individual than does one who hates the second individual; at the same time, a person who hates a second will require more kindness to changeover to love than does one who already cares for the second. Simply then, movement coming from one direction precipitates a catastrophe in a different place than does movement from a different direction, and this lag is called hysteresis.

The above is not intended to be a complete examination of catastrophe theory. What has been presented is a review of one catastrophe model to help explain the application of the theory to two consumer behavior examples.

Application of the Cusp Catastrophe to Two Consumer Behavior Situations

Brief reflection reveals that the cusp catastrophe is an alternative model for several areas of study in consumer behavior. The remainder of the paper is devoted to our interpretations of two such examples. Our intent is purely illustrative, consequently, although a wealth of research and thought may exist for both of our examples, we will not endeavor to review or critique. Also, our comments will be restricted to the cusp catastrophe model. Hopefully, those who are working in the areas of our two examples will envision the value of the conceptualization and be inspired to apply it in more detailed fashion.

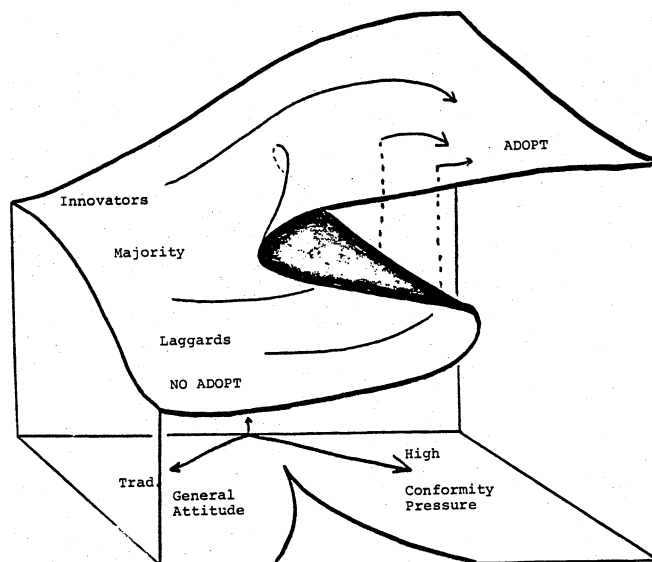
Example 1: Consumer Adoption

As a model of individual behavior, the cusp catastrophe offers an alternative view of the adoption decision. Its five conditions are satisfied in the following manner. Divergence exists in the form of the adopt/ no adopt decision facing the consumer. The phase of

evaluation prior to the adopt/no adopt decision satisfies the bimodality condition while inaccessibility is evident in that a consumer will resist ambivalence or half-adoption. Sudden transitions in behavior occur at the point of adoption, especially for later adopters. Finally, the hysteresis condition is evidenced by the fact that the decisions to adopt or to unadopt are unlikely to occur under the same conditions for all consumers.

To further illustrate adoption as a cusp catastrophe, Figure 4 represents the process as adoption/no adoption behavior overlaid on two control surfaces. In this interpretation, the control dimensions are: (1) conformity pressure as perhaps measured by the urgings of others or market penetration and (2) individual predispositions such as desire for novelty, recognition, acceptance, or some other relatively stable personal attribute. These are subsumed into a general attitude position ranging from liberal (acceptance of change) through moderate (deliberative and skeptical of change) to traditional (resistant to change), reminiscent of characteristics of various adopter groups. Seen this way, the cusp catastrophe model can simultaneously describe all classes of adopters (and non-adopters) across the time dimension, and it affords insight to the differences in their adoption decisions.

FIGURE 4
Cusp Catastrophe Interpretation of Consumer Adoption



Innovators progress through the adoption process by traversing the smooth surface of the cusp catastrophe defined by liberal attitudes and introduction of the innovation to the market. Their behavior is non-catastrophic in that they do not cross the bifurcation set. Adoption of new behavior is natural and uninhibited; moreover, the decision to stop adoption is a smooth transition back during this early stage. The majority of the population, however, faces a very different circumstance, for their moderate attitudes hold them at a state of no adoption until conformity pressure increases to the point at which they must evaluate the adoption decision. In the terminology of catastrophe theory, they are in the domain of the bifurcation set. Those with somewhat less moderate attitudes (e.g., the early majority) reach the edge of the no adoption behavior surface pleat and drop (up) to the adoption behavior surface. A snowball effect is seen as large segments of

the population adopt, increase conformity pressure, and cause (more skeptical) others to adopt. Ultimately, the innovation is adopted by all but a very few members of the population. At this point, the laggards or traditionalists occupy the behavior surface where the pleat is widest. Divergence for this group is greater than for any other adopter group; consequently, the change in behavior for them is literally greater than for anyone else. For this reason, they will resist change until they are forced to adopt.

Example 2: Consumer Complaining Behavior

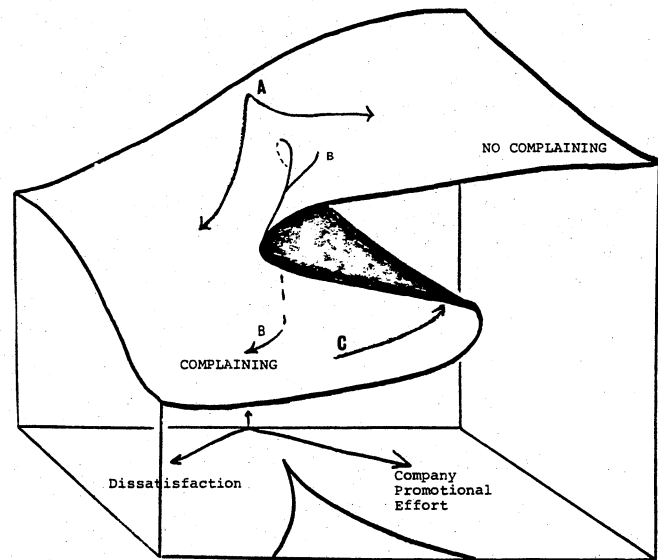
Our illustration of the adoption cusp catastrophe utilizes a dynamic control dimension (conformity pressure) matched against a stable psychological state. The model easily handles two dynamic control dimensions, as will be demonstrated in this example. Consumer complaining behavior has received an increasing amount of attention in recent years although no widely accepted models exist in our knowledge. In our present interpretation, consumer complaining behavior can be seen as a function of consumer dissatisfaction and company promotional efforts. Dissatisfaction for the individual consumer grows and declines with product use, changes in expectations, and/or information acquisition. Company promotional efforts exist in the magnitude and quality of persuasion, education, demonstration or other such efforts. The behavior surface is comprised of complaining or not complaining (divergence). Under the same circumstances some individuals complain while others do not (bimodality); sudden changes in consumers from noncomplainers to complainers and vice versa have been observed, and complaining behavior is dependent on prior conditions (hysteresis).

Viewed as the cusp catastrophe in Figure 5, complaining behavior and even the intensity of complaints becomes clear. When a consumer's experiences engender dissatisfaction with a brand or company, he will traverse the behavior surface to the complaint mode. However, it is likely that company promotional efforts will counteract dissatisfaction to some degree early in the process. Perhaps as a consequence of personal style, some consumers will move out on the behavior surface and enter the bifurcation set on the no complain surface while others will choose the complaint surface (A). These two groups of consumers in the region outlined by the bifurcation set are undoubtedly the firm's marginal customers. One group persists in mild complaints to dealers, salesmen, or even company executives. The other group does not complain, but undoubtedly will welcome a competitor who will respond to their needs more effectively. Given that there are no satisfactory alternatives available, the situation can easily become escalated. That is, greater dissatisfaction as a result of greater product usage will draw the noncomplainers closer to the edge of complaining behavior (B). It will simultaneously push the mild complainers into a more militant posture. The reaction of the company to these symptoms is quite natural: it increases its promotional activities in an attempt to counteract the situation. Seen as a cusp catastrophe, however, the net result is to draw the consumers further out on the pleat and to thereby increase the potential of intensive complaining behavior.

At this point, the company has a tiger by the tail, even though it may not realize the situation. It cannot reduce its promotional activities; nor can it back off from its previous claims, for the model predicts an immediate and catastrophic consequence in the form of a flood of complaints to the company and probably consumer protection and other regulatory agencies as well. Moreover, when consumers change from noncomplainers to complainers under these circumstances, hysteresis dictates that massive amounts of company promotion will be

required to restore them to a noncomplaining attitude (C).

FIGURE 5
Cusp Catastrophe Interpretation of
Consumer Complaining Behavior



Fortunately, however, the model affords a prescription for the company. It must reduce dissatisfaction through product modification or restoration of damages. Actions, not words, will return consumers to the noncomplaining equilibrium state by reducing their dissatisfaction.

Conclusion

These two examples are illustrative of the ability of one catastrophe model to describe consumer behavior. It should be clear that while mathematically complex, the conceptual aspects of catastrophe theory are easy to comprehend and are sufficiently flexible to serve our needs. Given that consumer behavior is multidimensional and subject to abrupt changes, it will be to our advantage to scrutinize the applicability of the cusp and other catastrophe models in the near future.

References

- B. H. Arnold, *Intuitive Concepts in Elementary Topology* (Englewood Cliffs, New Jersey: Prentice-Hall, 1962).
- Stephen Barr, *Experiments in Topology* (New York: Thomas Y. Crowell Co., 1964).
- T. Oliva and C. Capdevielle, "Catastrophe Theory: A Collective Bargaining Application," Forthcoming in the *Proceedings of the National Academy of Management Annual Meetings*, (August, 1977).
- L. Starobin, "Our Changing Evolution: Strategies for 1980," in *General Systems XXI* (Washington, D.C.: Society for General System Research, 1976) 3-46.
- R. Thom, *Stabilite Structurale et Morphogenese* (New York: Buyamin, 1972).
- E. C. Zeeman, "Catastrophe Theory," *Scientific American* (April 1976), 65-83.

TESTING COMPETING MODELS OF CONSUMER DECISION
MAKING IN THE PREVENTIVE HEALTH CARE MARKETPLACE

Richard L. Oliver, University of Iowa
Philip K. Berger, University of Kentucky

Abstract

Three models borrowed from social psychology, medical sociology, and communication theory which have been proposed to predict behavioral intention were compared on the basis of their conceptual strength and predictive accuracy in a study of a community swine flu vaccination program. The results suggest that models incorporating evaluative components, normative influences, and intervening summary concepts may yield a greater understanding of health care decisions. Implications of the findings were discussed.

Introduction

The application of behavioral concepts to health product and service marketing was a natural outgrowth of the broadened marketing concept (Kotler and Levy, 1969). Early attempts to construct comprehensive models (Zaltman and Vertinsky, 1971; Becker, 1974; Rosenstock, 1974) were followed by various applications in medical sociology (Cf. Becker and Maiman, 1975, for review) and a recent symposium on behavior in the health marketplace (Newman, 1976). Unfortunately, consumer behaviorists have been slow to apply the emerging knowledge gleaned from studies on non-health related products to the medical field.

The purpose of this paper is to compare the extent to which a model currently used by medical sociologists to explain preventative health behavior and a well accepted behavioral intention model used in the marketing community can predict intentions to receive the swine flu vaccination. In addition, a recent hybrid model focusing on drive arousal and reduction tendencies will also be contrasted with the other two models. Hopefully, these comparisons will permit medical care and public health researchers to modify their present perceptions of preventative health and health compliance behavior to include conceptual modifications suggested in the social-psychological and consumer behavior areas. Moreover, it will also be shown that many of the concepts used in health behavior models are isomorphic with the more traditional views of market behavior.

The Models

The three models to be tested include Fishbein's (Fishbein and Ajzen, 1975) Behavioral Intention Model (BIM), the Health Belief Model (HBM) as discussed by Becker and Maiman (1975), and a drive component model proposed by Wortzel (1976) to predict health care decisions. All are to some degree "instrumentality" schemata as they suggest that action is undertaken because it is perceived to result in or block various life outcomes. Each model is specified in some detail in the following sections.

The Behavioral Intention Model. Because consumer behaviorists will be most familiar with Fishbein's BIM,

This research was partially supported by grants from the University of Kentucky Graduate School and the Lexington-Fayette County Health Department.

only a brief discussion is presented here. Based on Dulany's (1968) theory of propositional control, Fishbein (1967; Fishbein and Ajzen, 1975) proposed that one's intention to perform a specific behavioral action is an additive function of two components, attitude toward the behavior and normative social influences. One's attitude toward the behavior, in turn, is a function of the sum of products of beliefs about the consequences of the behavior and one's evaluation of those consequences. In similar fashion, the summary perception of normative social influences, referred to as one's subjective norm, is believed to be a function of the sum of products of normative beliefs attributed to various referent persons and one's motivation to comply with those beliefs. Other stimulus conditions are thought to influence intention through the attitudinal and normative components.

The Health Belief Model. As summarized by Becker and Maiman (1975), the HBM predicts that people will engage in preventive health behavior if "they possess minimal levels of relevant health motivation and knowledge, perceive themselves as potentially vulnerable and the condition threatening, are convinced of the efficacy of intervention, and see few difficulties in undertaking recommended action (p. 12)." Unlike Fishbein's Behavioral Intention Model, the HBM is less formalized in its constituent components and the measurement of its variables. Based primarily upon articles by Becker (1974), Becker and Maiman (1975), Maiman and Becker (1974), and Rosenstock (1974), the version of the HBM discussed in the next section has been adopted for the purpose of this study.

Proponents of the HBM suggest that the probability of an individual engaging in a preventive health action is a function of three factors, a "benefits-barriers" analysis of the advantages and disadvantages of the health prevention activity, the "perceived threat" associated with the condition or illness, and various "cues to action" which include both mass media and interpersonal communications.¹ The perceived threat construct, in turn, is believed to have two antecedents, namely one's perceived susceptibility to the health problem and the anticipated severity of the problem if contracted. In a manner similar to that employed in the BIM model, other determinants are thought to affect the antecedents of one's intention rather than the intention itself.

In summary, an individual's perceptions of susceptibility to and severity of the focal health problem constitute the inherent threat posed by the disease. Given that a preventive health action is available, the model assumes that the person evaluates the action in terms of its perceived potential benefits to reduce this threat as well as the potential barriers or costs to taking the

¹The model described here departs in one important respect from earlier versions (e.g., Rosenstock, 1974). The cues to action component in prior formulations impacted directly on the perceived threat component and indirectly on the likelihood of behavior variable. A careful reading of the literature however, suggests that cues to action more appropriately affect the likelihood of behavior directly rather than through perceived threat. This latter interpretation is the one tested in this paper.

same action. Finally, the individual is also assumed to be further influenced by personal and impersonal recommendations. These three major components are believed to define on's subjective state of readiness to take preventive action.²

A Comparison of the BIM and HBM

A number of differences exist between Fishbein's BIM and the HBM. Some have been suggested by Jaccard (1975), others derive from exchanges between Songer-Nocks (1976) and Fishbein and Ajzen (1976) while some interpretations are novel to this paper. Each is summarized briefly in the following discussion.

First, Fishbein's BIM is designed to predict intention to act, not action itself. The original version of the HBM contains no variable similar to behavioral intention but is designed to predict actual behavior. If this difference is taken at face value, it would not be possible to compare the two models. For the purposes of this paper, however, the criterion for both models was identified as behavioral intention. If one is willing to assume a high positive correlation between intention and behavior, the data are useful for comparing the two models.³

Second, the BIM requires that beliefs concerning the consequences of the preventive health act be multiplicatively combined with one's evaluation of those consequences. In the HBM, evaluations of the consequences of the act are not explicitly called for and, as a result, may not be measured. Since the HBM is not as formally specified as the BIM, the degree of precision surrounding the explication of consequences is not known. It can be said, however, that the outcomes of taking action are assessed as perceived benefits and barriers. Moreover, it appears that all benefits are assumed to be universally evaluated as good while the barriers are assumed to be consistently evaluated as bad.

Third, social norms are a necessary component of Fishbein's model. In contrast, the analogous normative influences in the HBM are listed among many other variables in the cues to action construct. By the same line of reasoning, the BIM conceptualization does not explicitly provide for various mass media messages but rather subsumes them under the category of exogenous stimulus conditions. Thus, depending on the manner in which a particular investigator operationalizes influences external to the individual in either model, the relevant social norms or media impacts may not be included.

A fourth difference between the two models concerns the fact that the HBM maintains, as separate components, the perceived susceptibility, severity, and threat of the potential illness. Because Fishbein's model was developed in non-health related environments, the need to explicitly provide for emotional fear arousal variables was not evident. Consequently, the BIM is

²Some versions of the model suggest that demographic and personality variables affect perceptions of the major variables. These individual difference constructs have been purposely omitted in order to focus on the cognitive concepts in the HBM.

³The relationship between intention and behavior is an empirical issue which the authors intend to address at a future time when behavioral data collected by the local health department can be retrieved.

limited, to a certain extent, to the rational side of a preventive health decision. Although the HBM does not explicitly state the rules for combining susceptibility and severity to arrive at perceived threat, it does provide for their specification in the model.

To aid the reader in comparing the two instrumentality models described in the preceding discussion, the constructs in the BIM and HBM are juxtaposed on the basis of their isomorphic similarities in Table 1 along with the components of the Drive Reduction Model (DRM). Discussion of this latter model, due to its somewhat different orientation, has been reserved until this time.

Table 1

Conceptual Comparison of Three Models For the Prediction of Preventive Health Behavior Intentions

Construct	Behavioral Intention Model	Health Belief Model	Drive Reduction Model
Likelihood of Consequences	Yes	Yes (Benefits- Barriers)	Yes (Drive Reduction)
Evaluation of Consequences	Yes	No	No
Attitude Toward Behavior	Yes	No	No
Normative Beliefs	Yes	Yes (Cues to Action)	No
Motivations to Comply	Yes	No	No
Subjective Norm	Yes	No	No
Perceived Susceptibility	No	Yes	Yes } Drive Arousal
Perceived Severity	No	Yes	Yes
Perceived Threat	No	Yes	No
Exogenous Influences	No	Mass Media	No
Criterion	Behavioral Intention	Likelihood of Action Occurring	Motivation to Act

The Drive Reduction Model. A preventive health model based on the fear arousing properties of an illness and the capacity of a health action to reduce that fear has been proposed by Wortzel (1976) as adapted from earlier work by Bauer and Cox (1963) on the rational and emotional content of communication influences. Bauer and Cox originally suggested that the two critical dimensions which affect a consumer's motivation to engage in a given act are the level of emotional arousal resulting from failure to take action and the subjective probability that that same action will prevent the anticipated problem which is generating this arousal. The consumer's level of felt emotion is considered to be the drive arousal dimension while the subjective degree of success attached to a preventive action is believed to be the drive reduction mechanism.

The authors predict that, when the level of the subjective probability of success relative to arousal is sufficiently high, a person will be motivated to engage in the act; when that contrast goes below a certain (unstated) value, the person will choose not to engage in that behavior.

Wortzel (1976) later operationally redefined the Bauer and Cox (1963) drive arousal and drive reduction dimensions in term of the variables encompassed by the HBM. Specifically, the drive arousal dimension was conceptualized as the product of an individual's perception of his susceptibility to a disease and the perceived seriousness of the same disease. The drive reduction dimension, in turn, was defined as the difference between the subjective probability that a preventive health action will result in avoiding the disease and one's subjective assessment of the cost of preventive behavior. Predictions from this reoperation-alized model are identical to those of Bauer and Cox, i.e., action is a function of the level of drive reduction as related to drive arousal.

Table 1 shows how the variables suggested in the DRM compare to those discussed with regard to the BIM and HBM. Further elaboration is necessary, however, because a number of subtle differences exist in the variables used to construct the major concepts in the model.

First, the number of outcomes of the preventive behavior used in the DRM is somewhat restricted. Only one positive consequence, that of immunity from the disease, is specified while the negative outcomes are limited to "costs" which are subjectively identified. Thus, the DRM may exclude a number of positive outcomes (e.g., protecting one's family from contagion) or negative outcomes such as uncomfortable, unknown, side effects. In addition, the subject is not asked to make evaluations of the degree of immunity received or of the costs, a problem which is also inherent in the HBM.

Second, the DRM is more limited and also more explicit in its operational definition of perceived susceptibility to and perceived seriousness of the disease. While the HBM admits many diverse operational definitions of these variables (Becker and Maiman, 1975), the DRM requires a probabilistic judgment of susceptibility in the absence of the preventive health behavior. In an equally rigorous manner, seriousness is viewed as a combination of all ill effects of the disease. Also of note is the fact that the DRM specifies that the combinatorial rule for these two variables is multiplicative. Jaccard (1975) has noted that in various formulations of the HBM, both additive and multiplicative rules have been used.

Finally, the DRM appears to suffer from a greater degree of conceptual deficiency than either of the other two models. In addition to the previously discussed omissions, no measures of personal or mass media influences are specified nor is the summary construct of perceived threat. Rather, these variables appear to have been considered either redundant, exogenous, or uncontrollable.

Hypothesis

The degree to which the three models predict one's intention to receive a swine flu vaccination is the central focus of this paper. The corresponding null hypothesis specifies that no differences will be observed. The next section discusses the methodology employed to test this notion.

Method

Design

A cross-sectional field study of a community swine flu vaccination program was undertaken in October 1976. Questionnaires measuring beliefs, attitudes, and intentions concerning the swine flu inoculation campaign were mailed to residents one week before the vaccine first became available in that community. A number of other municipalities across the country had already initiated their vaccination programs.

Subjects

Two thousand residents of a medium-size midwestern community were selected from the telephone directory using a systematic random sampling procedure to receive the questionnaire used in the study. In addition, 1000 students from a major state university located in the community were also selected by the same random process from the student directory to participate in the survey.

Measures

Behavioral Intention Model. Fishbein (1975) requires that all salient consequences of the behavioral referent as well as the relevant reference persons be identified. This task was facilitated by the controversial nature of the swine flu inoculation campaign. Because of the widespread media coverage of the swine flu program, the pros and cons of immunization were fairly well documented (e.g., Boffey, 1976; Spivak, 1976). In addition, a culling of varied media presentations including local TV interviewing provided a sufficient list of consequences and referent persons and groups. The number of factors in each list was later reduced through pretesting to eight consequences of obtaining the swine flu shot, five consequences of failing to obtain the shot, and five social referents.⁴

To obtain a probabilistic measure of beliefs about the consequences, the respondents were asked to scale the possibility of occurrence of each consequence if they got (did not get) a swine flu shot on a five point scale ranging from "no chance" (0) through "50-50" (.5) to "certain" (1). The evaluative component of the model was measured by asking respondents to evaluate each consequence on a five point good-bad scale ranging from "very bad" (-2) to "very good" (+2). In a similar manner, subjects indicated their normative beliefs with regard to each social referent by responding to the question "Do the following people want you to get a swine flu shot?", on a five point scale ranging from "yes" (+2) to "no" (-2). Finally, motivation to comply was obtained from four point items asking the respondent: "Does (social referent's) opinion matter to you?" Scale points ranged from "yes" (3) to "no" (0).

Summary constructs tapping one's overall attitude toward getting a swine flu shot and one's overall subjective norm were measured on independent scales. The attitude

⁴ The consequences of obtaining the shot were a mild case of the flu, pain from the shot, immunity from the flu, protecting family and friends from the flu, being inconvenienced, side effects or allergic reactions, feeling more comfortable with people, and possible fatality due to the shot. The consequences of not getting the shot included catching swine flu and, if one caught the flu, pain and discomfort, cost of medicine and doctor's bills, lost time from one's job and/or school, and death. The social referents were one's doctor, family or spouse, local health department, company or boss, and friends and neighbors.

toward the act measure used in the study was a nine item semantic differential scale⁵ while the subjective norm variable was a one item summary measure suggested by Fishbein and Ajzen (1975, p. 314).

The Health Belief Model. Perceived susceptibility, severity, and threat were measured on five, eight, and four item Likert scales respectively. Items for these scales were suggested by Rosenstock's (1974) review of earlier studies and by media statements specific to the swine flu campaign.⁶

To obtain a measure of the benefits-barriers concept, scores from the probability of consequences scale in the BIM (i.e., the belief component) were used without the evaluative dimension. Thus, the average of the probabilities assigned to the five unfavorable outcomes (barriers) was subtracted from the mean of probabilities assigned to the three favorable outcomes (benefits) to arrive at the benefits-barriers score.

The cues to action variable was obtained in a section of the questionnaire separate from that measuring the normative influences in the BIM and included references to TV, radio, and the print media as well as one's doctor, family, and friends. Respondents were asked to specify the media and persons used as information sources and to indicate whether this information was "for swine flu shots" (+2) at the one extreme of a five item scale, to "against" flu shots (-2) at the other. The sum of these responses was used to operationalize the cues to action construct.

The Drive Reduction Model. The drive arousal component of the DRM was measured with reference to the action null set, that of failing to get a swine flu shot. According to Wortzel (1976), this variable is defined as the probability of contracting swine flu in the absence of an inoculation times the anticipated severity of the problems associated with the disease. To obtain these measures, subjects were asked to first indicate their chances of catching swine flu if they did not get a shot (perceived susceptibility) and then to estimate their chances of incurring four negative outcomes if they did get flu (perceived severity), all on five item scales ranging from "no chance" to "certain". The four outcome scores were averaged and then multiplied by the susceptibility probability.

Drive reduction has been defined by Wortzel (1976) as the subjective probability of success associated with the preventive action less an assessment of costs of that action. These measures were taken from the probability of consequences scales discussed in the BIM. The probability of success variable was operationalized as the likelihood assigned to the one outcome "immunity from swine flu" while the costs of action were estimated by the sum of probabilities attached to the negative consequences of getting a swine flu shot. Drive reduction was then operationalized as the probability of immunity minus the average of the sum of costs.

Criterion. To obtain a continuous measure of behavioral intention, subjects were asked to indicate the "chances in 10" that they would get a swine flu shot on an eleven point scale ranging from "no chance (0) to "certain" (1).

⁵The coefficient alpha reliability of this scale was .94.

⁶The scales yielded coefficient alpha reliabilities of .68, .72, and .70 respectively.

Analysis

Simple correlations were first calculated between the criterion and the antecedents suggested by the various models to test whether the predictor variables were significantly associated at the zero order level. The intention criterion was later regressed on all suggested antecedents simultaneously to determine if each variable would make a significant independent contribution to the explanation of variance in intention. Finally, the coefficients of determination generated by each model were compared in an effort to identify the model with the greatest predictive potential. These analyses were performed for the sample as a whole and then by student and resident subsamples to determine if differences in model component weights existed across the two subject groups.

Results

Of the 3000 mailed surveys, 252 student and 148 resident questionnaires were returned as undeliverable for various reasons, the most common of which was "moved, no forwarding address." Of the remaining 2600 surveys, 335 usable student replies and 515 usable resident returns were received yielding response rates of 45% and 28% respectively. Of these respondents, twelve students and 46 residents indicated that they had already received the shot and were eliminated from the study resulting in a final sample size of 792 including 323 students and 469 residents. Fifty-one percent of the student respondents and 53% of the residents in the sample were male.

Correlations between the intention criterion and all variables suggested by the three models are shown in Table 2 for the total sample and the two subsamples. It is evident that all antecedent variables were significantly correlated with the criterion. Moreover, the magnitudes of these associations were higher in the BIM than in the other two models suggesting that the Fishbein approach may have greater validity. It is also interesting to note that, in those models where concepts are linked sequentially (the BIM and the HBM), the variables posited as having a direct effect on intention produced higher correlations than those suggested as having indirect effects only, a finding which supports the conceptual structure of these models.

Table 2 also shows that marked differences across subsamples were observed only for the cues to action construct in the HBM and for the drive reduction construct in the DRM. The data suggest that intentions of the community residents were more highly related to media and word-of-mouth influences than were students' intentions and that the drive reducing properties of the inoculation were somewhat more instrumental for resident intentions than for students. For the most part, however, the correlations obtained in the two samples for all models were surprisingly similar.

When the criterion was regressed on all variables in each model simultaneously, the results shown in Table 3 were obtained. It is immediately apparent that the BIM yielded coefficients of determination much larger than those obtained with the HBM which, in turn, produced higher coefficients than the DRM. Note further that all variables hypothesized to directly affect the intention criterion generated significant coefficients while those which are posited to have indirect effects (attitudinal and normative beliefs in the BIM; perceived susceptibility and severity in the HBM) resulted in attenuated coefficients when compared to the zero-order correlations in Table 2. This finding lends support to the sequential or intervening variable influences suggested by the BIM and the HBM.

Table 2

Correlations Between the Intention Criterion and Antecedents Suggested by Three Preventive Health Models

Model and Antecedent	Sample		
	Total	Student	Resident
Behavioral Intention Model			
Σ (Beliefs x Evaluations)	.46*	.43*	.47*
Σ (Norm. Beliefs x Motivations)	.55*	.50*	.56*
Attitude Toward Behavior	.70*	.70*	.70*
Subjective Norm	.60*	.59*	.62*
Health Belief Model			
Perceived Susceptibility	.18*	.18*	.18*
Perceived Severity	.31*	.29*	.31*
Perceived Threat	.39*	.38*	.38*
Benefits - Barriers	.42*	.38*	.44*
Cues to Action	.39*	.32*	.42*
Drive Reduction Model			
Drive Arousal	.32*	.34*	.31*
Drive Reduction	.34*	.28*	.37*

* p < .01

Table 3

Standardized Coefficients Obtained When Intention Was Regressed on All Suggested Antecedents: Three Models

Model and Antecedent	Sample		
	Total	Student	Resident
Behavioral Intention Model			
Σ (Beliefs x Evaluations)	.07*	.01	.10*
Σ (Norm. Beliefs x Motivations)	.06	.08	.04
Attitude Toward Behavior	.50**	.53**	.48**
Subjective Norm	.21**	.19**	.23**
R ²	.54**	.52**	.55**
Health Belief Model			
Perceived Susceptibility	.02	.00	.04
Perceived Severity	.07	.12*	.03
Perceived Threat	.21**	.24**	.18**
Benefits - Barriers	.29**	.27**	.32**
Cues to Action	.21**	.17**	.24**
R ²	.32**	.29**	.35**
Drive Reduction Model			
Drive Arousal	.30**	.31**	.29**
Drive Reduction	.31**	.23**	.36**
R ²	.21**	.17**	.24**

* p < .05

** p < .01

Table 3 shows that somewhat higher coefficients of determination were obtained when using the resident sample data. This finding was particularly evident for the HBM and DRM models and may arise from the higher zero order correlations for certain of the variables. It would be speculative, however, to attribute meaningful differences in model structure across samples to the small differences observed in the regression coefficients. Such differences may be more easily explained by multicollinearity effects or range differences in the variables.

Discussion

The results have shown that the BIM yielded much higher coefficients of determination than either of the other competitive models. As the reader may recall, Fishbein's conceptual framework differed in a number of important respects from other schemata and this may explain its predictive superiority. First, the BIM incorporates evaluative components at both the individual consequence level and at the summary attitude toward (evaluation of) the act level. This would appear to be critical because individual beliefs have little meaning unless one's affect toward a consequence or behavior is gauged (Fishbein and Ajzen, 1975). Thus, the implicit assumption in the HBM that all benefits are evaluated as equally good and all barriers or costs as equally bad is not necessarily accurate.

Second, the BIM explicitly includes the influences of referent persons and, in doing so, acknowledges that relevant others are instrumental in the formation of one's behavioral intentions. While the HBM does have a cues to action component which is somewhat isomorphic with Fishbein's normative beliefs, it does not provide for the various motivations to comply with those beliefs (which can be seen as another form of evaluation). Again, it may be inappropriate for proponents of the HBM to assume that all cues to action which support one's intention to engage in the preventive behavior are viewed as equally salient or credible by the consumer.

Finally, Fishbein's model includes two important intervening variables, attitude toward the act and subjective norm, between the summed attitudinal and normative beliefs and intention. These mediating constructs serve to "capture" the essence of the underlying beliefs and, if the results of this study are representative, suggest that individuals do, in fact, form summary judgments of positive and negative consequences and positive and negative social influences. These summary judgments, then, become the most direct influences (and therefore the best predictors) of intention. Neither the HBM nor the DRM contains these variables.

The difference in the coefficients of determination between the HBM over that of the DRM may be attributed to the greater number of variables in the former model. Specifically, the HBM explicitly recognizes external information sources on one's intention to engage in a behavior and is not as restrictive in terms of the benefits or favorable outcomes which can be considered in the model. Thus, proponents of drive models of preventive behavior may wish to add social and other influences in an effort to increase the explained variance in their criteria.

In summary, it appears that the inclusion of evaluative components, normative influences, and intervening summary measures enhance the degree to which a model aids in the explanation of preventive health intentions. The next section suggests some implications of these findings.

Implications

Profit and non-profit preventive health care marketers may find a number of the findings reported here of interest in improving their understanding of consumer health decisions. First, it is noteworthy that all three models investigated explained a statistically significant portion of the variance in behavioral intention suggesting that some of the antecedents of health intentions may be known. Consequently, practitioners and field researchers may wish to borrow from certain of the models to augment the decision making processes used by their organization.

Second, researchers concerned with the most complete determination of health care decisions (e.g., when an epidemic is imminent) may wish to combine concepts from two or more of the models. To illustrate, it appears that inclusion of one or more emotional components (e.g., perceived threat) in the BIM may increase its predictability beyond that obtained with the more rational components proposed by Fishbein.⁷ Alternatively, researchers with an HBM orientation may wish to include evaluations of the benefits and barriers as well as the individual's overall evaluation of the preventive health behavior.

Finally, it is clear that some variables in the models have more immediate impacts on one's intention decision. It follows that if a researcher or practitioner were interested primarily in prediction, a more parsimonious model consisting of direct effects only may suffice. In a similar manner, if one wished to attempt to change an existing behavioral intention, it may be more productive to work on the more immediate antecedents. For example, a person's overall subjective norm may be more amenable to influence in the BIM than would the orientations of the individual referent persons.

Limitations

The conclusions drawn here are contingent on the methodology employed to test the models. In view of this, certain readers may wish to question the authors' interpretation of model structure as well as the operationalization procedures used. To the extent that these criticisms are valid, the findings remain tentative. We would hope that others would take issue with this presentation and test alternative models and methodologies. Hopefully, a discourse of this nature will result in beneficial refinements and a more complete understanding of preventive health care decisions.

References

Raymond A. Bauer and Donald F. Cox, "Rational Versus Emotional Communications: A New Approach," in L. Arons and M. May, eds., Television and Human Behavior (New York: Appleton-Century-Crofts, 1963).

⁷ Further analysis of the data used here provides only tenuous evidence for this suggestion. When intention was regressed on the four Fishbein components and perceived susceptibility, severity, and threat simultaneously, the coefficients of determination increased by 2% for the total sample, by 4% for the student sample, and 2% for the resident sample. This is not to say, however, that greater improvements would not be obtained in other research settings.

Marshall H. Becker, "The Health Belief Model and Sick Role Behavior," Health Education Monographs, 2 (Winter, 1974), 409-19.

Marshall H. Becker and Lois A. Maiman, "Sociobehavioral Determinants of Compliance with Health and Medical Care Recommendations," Medical Care, 13 (January, 1975), 10-24.

Philip M. Boffey, "Anatomy of a Decision: How the Nation Declared War on Swine Flu," Science, 192 (May, 1976), 636-41.

Don E. Dulany, "Awareness, Rules, and Propositional Control: A Confrontation with S-R Behavior Theory," in D. Horton and T. Dixon, eds., Verbal Behavior and General Behavior Theory (New York: Prentice-Hall, 1968).

Martin Fishbein, "Attitude and the Prediction of Behavior," in Martin Fishbein, ed., Readings in Attitude Theory and Measurement (New York: Wiley, 1967).

Martin Fishbein and Icek Ajzen, Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research (Reading, Mass.: Addison-Wesley, 1975).

Martin Fishbein and Icek Ajzen, "Misconceptions about the Fishbein Model: Reflections on a Study by Songer-Nocks," Journal of Experimental Social Psychology, 12 (1976), 579-84.

James Jaccard, "A Theoretical Analysis of Selected Factors Important to Health Education Strategies," Health Education Monographs, 3 (Summer, 1975), 152-67.

Philip Kotler and Sidney J. Levy, "Broadening the Concept of Marketing," Journal of Marketing, 33 (January, 1969), 10-15.

Lois A. Maiman and Marshall H. Becker, "The Health Belief Model: Origins and Correlates in Psychological Theory," Health Education Monographs, 4 (Winter, 1974), 336-353.

Ian M. Newman, ed., Consumer Behavior in the Health Marketplace (Lincoln, Nebr.: Nebraska Center for Health Education, University of Nebraska, 1976).

Irwin M. Rosenstock, "The Health Belief Model and Preventive Health Behavior," Health Education Monographs, 4 (Winter, 1974), 354-86.

Elaine Songer-Nocks, "Situational Factors Affecting the Weighting of Predictor Components in the Fishbein Model," Journal of Experimental Social Psychology, 12 (January, 1976), 56-69.

Jonathan Spivak, "Program to Inoculate All Americans for Flu Appears to be Ailing," Wall Street Journal, 56 (June 18, 1976), 1ff.

Lawrence H. Wortzel, "The Behavior of the Health Care Consumer: A Selective Review," in Beverlee B. Anderson, ed., Advances in Consumer Research, 3 (1976), 295-301.

Gerald Zaltman and Ilan Vertinsky, "Health Service Marketing: A Suggested Model," Journal of Marketing, 35 (July, 1971), 19-27.

Michael J. Ryan, Columbia University

Abstract

An argument for expanding Fishbein's Behavioral Intention Model that posits non-additivity, the introduction of a new variable, and restructuring of the model by way of explicitly identifying components as antecedent, independent, and dependent variables is put forth. Evidence providing limited support is described and a modified expansion of the model is suggested as a guideline for future research.

Introduction

The popularity of applying and investigating the Fishbein Behavioral Intention or Extended Model within the context of consumer behavior is evidenced in a recent review (Ryan and Bonfield, 1975). More recently, efforts have focused on refinements or extensions of the original model (Bennett and Harrell, 1975; Lutz, 1975; Ahtola, 1975; Glassman and Fitzhenry, 1975). This research examines a portion of a causal chain proposed by Ryan and Bonfield (1975) as a framework for identifying relationships and specifying components as antecedent, independent, moderator, and dependent variables. The major focus regards non-additivity of attitudinal and social influences and the specification of an independent variable form of social influences.

The Fishbein Behavioral Intention Model

The Fishbein behavioral intention theory states that intention to perform a behavior in a given situation and subsequent performance of that behavior is a function of (1) attitude toward the performance of a specific behavior in a particular situation and (2) social norms influencing the individual's performance of the behavior. The central equation in the theory is commonly modeled as follows:

$$BI = (Aact) \omega_0 + \left(\sum_{j=1}^k NB_j MC_j \right) \omega_1$$

Where: BI = behavioral intention

$$Aact = \sum_{i=1}^n B_i a_i \quad (\text{attitude toward the performance of the act})$$

B_i = the expectation, i.e., the probability or improbability that the performance of a behavior will lead to an i th outcome

a_i = the positive or negative evaluation of the i th outcome

n = the number of salient outcomes

NB_j = the expectation, i.e., the probability or improbability, that the performance of the act is expected by a j th referent

MC_j = the motivation to comply or not to comply with the expectation of the j th referent

k = the number of salient referents

ω_0, ω_1 = empirically determined standardized regression coefficients (beta weights)

In previous research, the operational model has taken the form of a standardized regression equation and the magnitude of the beta weights has in turn been interpreted as an indication of the relative strength of attitudinal and social influences.

By combining attitudes and social influences Fishbein attempts to overcome a probable reason for the traditional lack of empirical support for an attitude-behavior relationship. Namely, that perceived social group influence variables must be considered in conjunction with attitudes to predict and understand behavior (Triandis, 1965; Warner and DeFleur, 1969; Wicker, 1969). Although it may be useful to separate these two sources of influence in regard to understanding behavioral intention formation and change strategies, Fishbein does not provide a conceptual basis for the independence stated in the additive model. This is a crucial point since the use of beta weights to indicate the relative influence of the variables on behavioral intention rests upon the additivity assumption. Many studies (Ajzen and Fishbein, 1973; Ryan and Bonfield, 1975) have produced statistically significant ω_0 and ω_1 beta weights thus yielding empirical support for the additivity notion. However, other theorists (Banton, 1965; Krech, *et al.*, 1962; Karlins and Abelson, 1970) have argued on a conceptual basis and furnished empirical evidence (Janis and King, 1954; King and Janis, 1956; Siegal and Siegal, 1957; Newcomb, 1958; Sherif, 1958; and Janis and Gilmore, 1965) indicating that attitude and social influences are related. Although these studies did not use expectancy-value models and conceptual differences make their total assessment difficult, this evidence together with the lack of a conceptual foundation, suggests that the additivity notion be questioned.

A Reconceptualization

The research carried out by Fishbein and his associates has usually incorporated a measure of Aact as the attitudinal component instead of $\sum B_i a_i$. Since these studies (Ajzen and Fishbein, 1973) were primarily concerned with prediction of behavior and behavioral intentions, this seems appropriate from an operational standpoint. In all cases the measurement of Aact was patterned after the multi-item, bipolar adjective, evaluative dimension of the semantic differential (Osgood, Suci, and Tannenbaum, 1957; Fishbein and Raven, 1962). This technique seems appropriate for predictive purposes since it is generally considered to produce reliable and valid measures of attitude. However, the bipolar adjective set does not have much explanatory power since it merely indicates the strength and direction of attitude. It does not reveal the elements comprising the attitudinal structure and therefore does not furnish diagnostic power useful in examining attitude formation and change. In those cases where predictive power or variable associations are the sole interest this is appropriate since a determination of structural elements is not necessary, but a reliable and valid measurement is essential. Given a choice, Aact is also easier to operationalize than $\sum B_i a_i$ since there is no need to determine salient outcomes, encounter problems arising from an erroneous or incomplete list, or run the risk that correlations with multiplied non-ratio scaled scores may be invariant (Schmidt, 1973).

The major thrust of Fishbein's work has been to conceptualize the attitude formation process as the summation of beliefs times evaluations (Fishbein, 1963).

The relationship between Aact and ΣB_{1a_1} has been supported in a number of studies (Ajzen and Fishbein, 1970; Bruce, 1971; Bright and Stummers, 1971; Jaccard and Davidson, 1972; Ajzen and Fishbein, 1972) and the interaction of B_1 and a_1 manipulations has produced changes in Aact (Lutz, 1975b). Thus, ΣB_{1a_1} appears appropriate for diagnosing Aact whereas Aact seems more appropriate than ΣB_{1a_1} for predicting behavioral intention. Empirical support for this notion is found in evidence provided by Lutz (1973). When he replaced Aact scores with ΣB_{1a_1} scores the multiple correlation on BI dropped from .647 to .240 and the attitudinal beta weight dropped from .646 to .238. This evidence was strengthened when Lutz (1975a) also observed changes in BI following changes in Aact that were induced by manipulating B_1 and a_1 values. In a pilot study for the present research it was found that ΣB_{1a_1} -BI correlations disappeared when Aact was partialled out (Ryan, 1974).

Although Fishbein does not use the term, this evidence indicates that ΣB_{1a_1} fits the use prescribed to antecedent variables which in this case is attempting to discover a first cause by clarifying the influences which precede the Aact-BI relationship. Thus, ΣB_{1a_1} is recast as an antecedent and Aact as an independent variable.

Following the same reasoning used to describe the two forms of attitude as antecedent and independent variables, the formulation of the normative component appears incomplete. The ΣNB_jMC_j formulation fits the use ascribed to explanatory variables but an independent form of social influence has not been specified that avoids operational problems and is more appropriate for prediction. In fact, within the context of Fishbein's theory the present understanding of the normative component is limited (Ajzen and Fishbein, 1973; Fishbein, 1975). Consequently, it seemed useful to look elsewhere for a framework within which to develop an independent form of the normative variable. Kelman's (1961) three processes of social influence; compliance, identification, and internalization, were used for this purpose.

Compliance occurs when an individual accepts influence from another person or group because he believes it will result in a satisfying social effect which may be the attainment of rewards or avoidance of punishments under the other's control. Identification occurs when an individual adopts the role of another. It differs from compliance in that the individual is not primarily motivated to please the other but to meet the other's own role expectations since they fit the individual's self image. Internalization occurs when an individual accepts influence because it is congruent with his value system. The reward is the content of the induced behavior.

Since internalization and identification involve rewards internal to the individual, these sources of social influence seem related to attitude. On the other hand, compliance involves an external reward since the individual assumes an actor's role, that is, a role not necessarily congruent with his internal values, in order to obtain an expected reward. Consequently, compliance may well be independent of attitude. However, Kelman holds that all three processes are present and interact in a given social situation and that the nature of the situation determines which particular process predominates and the type of interaction among them. Thus, some relationship between attitude and social compliance is expected with the strength of the relationship determined by the degree to which internalization and identification are present. In the intentions model such effects would be manifested in Aact whereas the social compliance dimension by itself, hereafter referred to as SC, seems independent from Aact. Thus, SC is posited as an independent variable form of social influence that is separable yet, in contrast to the additivity notion, related to Aact.

The notion of a separable influence variable based on external rewards also has some parallel with the development of Dulany's (1968) verbal response model from which Fishbein borrowed heavily in developing the Intention's Model. Dulany manipulated his variables in laboratory experiments that employed negative and positive reinforcers administered by an experimenter after subjects verbally responded to statements that were regarded as neutral in terms of the respondent's value system.

In summary, this reconceptualization views BI as a dependent variable, Aact and SC as independent variables related to each other, and ΣB_{1a_1} and ΣNB_jMC_j as antecedent variables. Subsequent to the initial work on this reconceptualization (Ryan, 1974), Fishbein and Ajzen (1975, pp. 16, 334, 407) have put forth a similar restructured model with three primary differences. First, attitudinal and social variables continue to be viewed as independently affecting BI. Second, although positing the same general flow of effects, identification of antecedent-independent variable relationships is not made explicit. Third, a new form of social influence is introduced as a generalized social norm (SN) identified as the person's perception that most people who are important to him think he should or should not perform the behavior in question (Fishbein and Ajzen, 1975, p. 302). Evidence supporting an association between ΣNB_jMC_j and SN and between SN and BI has been furnished (Glassman and Fitzhenry, 1975) although SN has yet to be shown as separable or independent from Aact. Fishbein (1975) introduced the importance concept based on an assumption that individuals would have little or no motivation to comply with unimportant others. This notion raises questions about both the separability of SN from Aact and the need to include MC in the antecedent variable. First, importance is likely to be based on value judgments or emotional feelings thereby indicating that "important others" would be internalized and reflected in Aact. Second, if importance guarantees a strong motivation to comply, this removes the necessity for including MC in the normative structure thereby rendering the theory less general since it would exclude motivations not to comply with a perceived social influence. For example, a referent expectation would be negatively related to behavioral intention if the individual was motivated not to comply with a positive influence. In the case of social compliance, such a situation may arise if the individual places little or no value on the expected reward from the referent as a consequence of performing the behavior. This argument parallels the thinking behind the original Fishbein (1967) expectancy times evaluation approach to attitude.

Research Objectives

This research tested the proposed associations among ΣB_{1a_1} , ΣNB_jMC_j , Aact, SC, and BI. It was not possible to empirically compare SC and SN since the research was executed prior to the appearance of SN. In addition to the value of replication, attitudinal variable relationships were included since it seemed unrealistic to test social influence variables without including their relationships with all other variables in the system. More specifically, support was sought for the following relationships as specified by the reconceptualization.

In terms of the attitudinal variables, attitude toward the act (Aact) should be directly related to its underlying structure (ΣB_{1a_1}) and behavioral intention (BI) should be directly related to attitude toward the act (Aact). Behavioral intention (BI) should be indirectly associated with attitudinal underlying structure (ΣB_{1a_1}). That is, behavioral intention (BI) should be more strongly associated with attitude toward the act (Aact) than it is with attitudinal underlying structure (ΣB_{1a_1}).

In terms of social influence variables, social compliance (SC) should be directly related to its underlying

significant correlations between their respective measurement scores. Indirectly linked variable correlations should lower in magnitude and become statistically insignificant when the moderator variable score is partialled out. Variables posited to have no relationship should produce statistically insignificant correlations between their respective measurement scores.

Further testing of the antecedent-independent variable designation involved an examination of beta weight and R^2 values when the respective variables were placed in the central equation. In order to support this notion the model employing Aact and SC should produce larger R^2 s than those produced using ΣB_{iaj} and ΣNB_jMC_j as predictor variables. In addition, it was expected that the Aact attitudinal beta weights would predominate for both brands and that the SC weight would be larger in the Ultra Brite versus the Crest model. Variations from these beta weight expectations when using ΣB_{iaj} and ΣNB_jMC_j as predictors would indicate that these variables are not appropriate independent variable forms.

Findings and Discussion

The simple and partial correlations used to test the attitudinal and normative variable associations are presented in Table 2. Correlation Sets 1 through 4 test attitudinal variable associations. The positive, statistically significant correlations obtained for both brands in Sets 1 and 2 support the direct associations between the attitudinal variables ΣB_{iaj} -Aact and Aact-BI. The statistically significant correlations in Set 3 for the indirectly linked variables ΣB_{iaj} -BI became statistically insignificant (Set 4) when Aact was partialled out. This evidence supports the associations posited among the attitudinal variables.

Correlation Sets 5 through 8 tested normative variable associations. The positive, statistically significant correlations for both brands in Sets 5 and 6 support the direct associations between the social influence variables ΣNB_jMC_j -SC and SC-BI. The statistically significant correlations in Set 7 for the indirectly linked variables ΣNB_jMC_j -BI became statistically insignificant

TABLE 2
Simple and Partial Correlations

Set	Brand	Variables	r
1.	Ultra Brite	ΣB_{iaj} -Aact	.368**
	Crest	ΣB_{iaj} -Aact	.633**
2.	Ultra Brite	Aact-BI	.653**
	Crest	Aact-BI	.635**
3.	Ultra Brite	BI- ΣB_{iaj}	.247**
	Crest	BI- ΣB_{iaj}	.317**
4.	Ultra Brite	BI- ΣB_{iaj} ·Aact	.010
	Crest	BI- ΣB_{iaj} ·Aact	-.142
5.	Ultra Brite	ΣNB_jMC_j ·SC	.559**
	Crest	ΣNB_jMC_j ·SC	.527**
6.	Ultra Brite	BI-SC	.562**
	Crest	BI-SC	.368**
7.	Ultra Brite	BI- ΣNB_jMC_j	.388**
	Crest	BI- ΣNB_jMC_j	.462**
8.	Ultra Brite	BI- ΣNB_jMC_j ·SC	.108
	Crest	BI- ΣNB_jMC_j ·SC	.339**

* $p < .05$

** $p < .01$

when SC was partialled out for Ultra Brite but remained significant for Crest (Set 8). The Crest partial correlation was smaller than its indirect correlation (Set 7) indicating that the removal of SC variance did lower the BI- ΣNB_jMC_j relationship. These findings largely support the proposed associations among the social influence variables. However, the failure of the statistical test to indicate that the entire Crest indirect association was due to the moderating effects of SC mitigates these findings and renders the support tentative.

Correlations testing the relationships among social influence and attitudinal variables are presented in Table 3. Correlation Sets 1 and 4 tested hypothesized

TABLE 3
Simple Correlations

Set	Variables	Ultra Brite	Crest
		r	r
1.	ΣB_{iaj} - ΣNB_jMC_j	.113	.312**
2.	ΣB_{iaj} -SC	.283**	.358**
3.	ΣNB_jMC_j ·Aact	.546**	.623**
4.	Aact·SC	.689**	.581**

* $p < .05$

** $p < .01$

associations. The positive statistically significant correlations in Set 4 support the associations between the independent variables Aact and SC. However, in Set 1 only the Crest correlation between the antecedent variables is significant and its value indicates the strength of the relationship was only moderate. This finding is unexpected, especially in view of the strong support for the independent variable relationship. It may be that the correlation of the two variable scores, each computed through the multiplication of non-ratio variable scores, may have attenuated these correlations.

Correlation Sets 2 and 3 tested hypothesized lack of associations among variables. The positive statistically significant correlations in both sets provide evidence disconfirming this hypothesis. In retrospect these findings are not surprising given the demonstrated colinearity between Aact and SC and the previous discussion concerning internalization. If certain specific sources of social influence are internalized it seems logical that these sources would become a part of attitude. This could happen, for example, if the focal person assumes a relevant other's expectation is congruent with his feelings or if a commonality of information about the product exists between the focal person and relevant others.

The findings concerning the use of the posited independent and antecedent variables in the central equation are shown in Table 4. The R^2 s were stronger for both brands and the social influence beta weight values were congruent with expectations when Aact and SC were used as independent variables whereas the weights were incongruent with expectations when ΣB_{iaj} and ΣNB_jMC_j were used as predictors. These findings lend additional support to the notion that Aact and SC are more appropriate for use as independent variables than are ΣB_{iaj} and ΣNB_jMC_j . In spite of these findings, however, the independent variable colinearity suggests that the

TABLE 4
Beta Weights and R² Estimates

$$BI = (\sum_{i=1}^n B_{ia_i})w_0 + (\sum_{j=1}^k NB_jMC_j)w_1$$

	Beta Weights		R ²
	$\sum B_{ia_i}$	$\sum NB_jMC_j$	
Crest	.19*	.40**	.25
Ultra Brite	.37**	.21*	.19

$$BI = (Aact)w_0 + (SC)w_1$$

	Beta Weights		R ²
	Aact	SC	
Crest	.64**	.01	.41***
Ultra Brite	.51**	.21*	.45

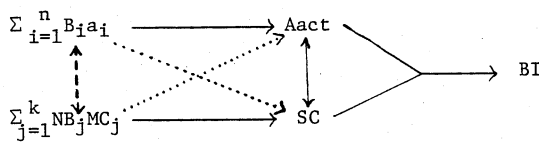
* p < .05
** p < .01
*** Due to statistical insignificance of the Crest SC beta weight, R² is based on simple Aact-BI correlations.

central equation beta weight interpretations may not be straight forward. First, an assumption must be made that colinearity does not cause beta weight instability. Second, the joint contribution of the predictors (c.f., Ferber, 1949) to BI variance should be considered since this effect may be stronger than the direct effects indicated by the beta weight values.

Summary and Conclusions

The findings from this research are summarized diagrammatically in Figure 1. The clear cut findings concerning the attitudinal variables are congruent with a large

FIGURE 1
A Summary of the Findings



Solid lines represent hypothesized relationships that were supported. Dotted lines represent supported relationships that were not hypothesized. Dashed lines represent hypothesized relationships that were not supported.

body of research investigating the expectancy value approach to attitude and a smaller but growing body of knowledge regarding relationships among cognitive, attitudinal, and behavioral intention variables. However, the findings concerning social influence variables and crossover effects were not straightforward and cannot be viewed within a framework provided by a body of research. The tentative evidence reported here constitutes the only support for the SC variable. More evidence is needed regarding SC measurement validity and applications across other products, brands, and groups.

The latter suggestion is especially appropriate given the limited generalizability of the sample used in this research. Also, other independent forms of social influence have been hypothesized (Ahtola, 1975; Lutz, 1975c) in addition to SN. Empirical comparisons of these different approaches are yet to be made. Consequently, the specification of meaning for a social influence independent variable is not a settled issue.

Given the state of knowledge, the suggested relationships in Figure 1 concerning crossover and social influence variable relationships are best viewed as tentative suggestions to guide future research. It is hypothesized that the indirect paths between the antecedent and outcome variable, indicated by dotted and dashed lines, represent weaker relationships than the direct paths, indicated by solid lines. The associations in this direct and indirect network seem susceptible to empirical testing through the use of path analysis (Nygren, 1971). However, the examination of associations is considered only a starting point rather than a complete investigation of what is proposed as a possible causal chain. Additional testing of causal order and lack of spuriousness would follow before causality could be supported or disconfirmed (Popper, 1959). Causal ordering, for example, could be tested with experimental design procedures outlined by Perreault and Darden (1975) that would consider the antecedent variables as factors and attitude, social compliance, and behavioral intention as multiple criterion variables.

More evidence supporting these notions would lead to a reduced form model since there are many variables in addition to those included that have been hypothesized to explain buyer behavior (Andreason, 1965; Nicosia, 1966; Howard and Sheth, 1969). However, there are two major advantages to the reduced form model. First, the relationships among variables are explicitly structured. Second, the level of abstraction allows the development of operational definitions for all variables. Consequently, the model can be shown to fit or not fit empirical regularities, a requirement of all good theories (Kaplan, 1964).

References

O. T. Ahtola, "Toward a Vector Model of Intentions," in B. B. Anderson (Ed.), *Advances in Consumer Research*, Volume III (Cincinnati: The Association for Consumer Research, 1975), 469-76.

A. R. Andreason, "Attitudes and Customer Behavior: A Decision Model," in L. Preston (Ed.), *New Research in Marketing* (California: University of California at Berkeley, Institute for Business and Economic Research, 1965), 1-16.

I. Ajzen and M. Fishbein, "Attitudinal and Normative Variables as Predictors of Specific Behaviors," *Journal of Personality and Social Psychology*, 27(1973), 41-57.

M. Banton, *Roles: An Introduction to the Study of Social Relations* (New York: Basic Books, 1965).

P. D. Bennett and G. D. Harrell, "The Role of Confidence in Understanding and Predicting Buyers' Attitudes and Purchase Intentions," *Journal of Consumer Research*, 2(September, 1975), 110-117.

H. M. Blalock, Jr., "Theory Building and Causal Inferences," in H. M. Blalock and A. B. Blalock (Eds.), *Methodology in Social Research* (New York: McGraw-Hill, 1968).

- Francis S. Bourne, "Group Influence in Marketing and Public Relations," in R. Likert and S. P. Haynes (Eds.), Some Applications of Behavioral Research (Paris: UNESCO, 1957), 208-224.
- B. E. Bright and H. R. Stammers, "Advertising, Promotion, and Medical Detailing," Proceedings, Seminar on Pharmaceutical Research (Amsterdam: European Society of Market Research, 1971).
- J. Bruce, "First Experiences with Fishbein Theory and Survey Methods," Proceedings, Seminar on Translating Advanced Advertising Theories into Research Reality (Amsterdam: European Society of Market Research, 1971, 25-39).
- A. B. Cowling, "Use of Elicitation Technique for Producing Dimensions of Brand Choice," Proceedings, British Market Research Society, 1973, 139-156.
- L. J. Cronbach, "Coefficient Alpha and the Internal Structure of Tests," Psychometrika, 16(1951), 297-334.
- D. E. Dulany, "Awareness, Rules, and Propositional Control: A Confrontation with S-R Behavior Theory," in D. Horton and T. Dixon (Eds.), Verbal Behavioral and General Behavior Theory (New York: Prentice-Hall, 1968).
- R. Ferber, Market Research (New York: McGraw-Hill, 1949).
- M. Fishbein, "A Consideration of Beliefs, and Their Role in Attitude Measurement," in M. Fishbein (Ed.), Readings in Attitude Theory and Measurement (New York: Wiley, 1967), 257-266.
- M. Fishbein, "An Investigation of the Relationships Between Beliefs About an Object and the Attitude Toward that Object," Human Relations, 16(1963), 233-240.
- M. Fishbein, "Extending the Extended Model: Some Comments," in B. B. Anderson (Ed.), Advances in Consumer Research, Volume III (Cincinnati: The Association for Consumer Research, 1975), 491-497.
- M. Fishbein and I. Ajzen, Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research (Reading, Mass.: Addison-Wesley Publishing Co., 1975).
- M. Fishbein and B. H. Raven, "The AB Scales: An Operational Definition of Belief and Attitude," Human Relations, 13(1962), 35-44.
- M. Glassman and N. Fitzhenry, "Fishbein's Subjective Norm: Theoretical Considerations and Empirical Evidence," in B. B. Anderson (Ed.), Advances in Consumer Research, Volume III (Cincinnati: The Association for Consumer Research, 1975, 477-480).
- J. A. Howard and J. N. Sheth, The Theory of Buyer Behavior (New York: Wiley and Sons, 1964).
- J. J. Jaccard and A. R. Davidson, "Toward an Understanding of Family Planning Behaviors: An Initial Investigation," Journal of Applied Social Psychology, 2(1972), 228-235.
- I. L. Janis and J. B. Gilmore, "The Influence of Incentive Conditions on the Success of Role Playing Modifying Attitudes," Journal of Personality and Social Psychology 1(1965), 17-27.
- I. L. Janis and B. T. King, "The Influence of Role Playing on Opinion Change," Journal of Abnormal and Social Psychology, 49(1954), 211-218.
- M. Karlins and H. I. Abelson, Persuasion (New York: Springer, 1970).
- A. Kaplan, The Conduct of Inquiry (San Francisco: Chandler Publishing, 1964).
- H. C. Kelman, "Process of Opinion Change," Public Opinion Quarterly, 25(1961), 57-78.
- B. T. King and I. L. Janis, "Comparison of the Effectiveness of Improvised vs. Non-improvised Role Playing in Producing Opinion Changes," Human Relations, 9(1956), 177-186.
- D. Krech, R. S. Crutchfield and E. L. Ballachey, Individual in Society (New York: McGraw-Hill, 1962).
- R. J. Lutz, "An Experimental Investigation of Causal Relations Among Cognitions, Affect and Behavioral Intention," Working Paper No. 34, Center for Marketing Studies, University of California, Los Angeles (October, 1975a).
- R. J. Lutz, "Changing Brand Attitudes Through Modification of Cognitive Structure," The Journal of Consumer Research, 1(March 1975b), 49-59.
- R. J. Lutz, Cognitive Change and Attitude Change: A Validation Study. Unpublished doctoral dissertation, University of Illinois-Urbana, 1973.
- R. J. Lutz, "Conceptual and Operational Issues in the Extended Fishbein Model," in B. B. Anderson (Ed.), Advances in Consumer Research, Volume III (Cincinnati: The Association for Consumer Research, 1975c), 469-476.
- T. M. Newcomb, "Attitude Development as a Function of Reference Groups: The Bennington Study," in E. E. Macoby, et al. (Eds.), Readings in Social Psychology (3rd ed.) (New York: Holt, Rinehart and Winston, 1958).
- F. M. Nicosia, Consumer Decision Processes (Englewood Cliffs, New Jersey: Prentice-Hall, 1966).
- J. C. Nunnally, Psychometric Theory (New York: McGraw-Hill, 1967).
- G. T. Nygreen, "Interactive Path Analysis," The American Sociologist, 6(February, 1971), 3-37.
- C. E. Osgood, G. J. Suci and P. H. Tannenbaum, The Measurement of Meaning (Urbana: University of Illinois, 1957).
- William D. Perreault, Jr. and William R. Darden, "Unequal Cell Sizes in Marketing Experiments: Use of the General Linear Hypothesis," The Journal of Marketing Research (1975), 333-342.
- K. R. Popper, The Logic of Scientific Discovery (New York: Basic Books, 1959).
- M. J. Ryan, "An Empirical Test of a Predictive Model and Causal Chain Derived from Fishbein's Behavioral Intention Model and Applied to a Purchase Intention Situation," unpublished doctoral dissertation, Department of Business Administration, University of Kentucky, 1974.
- M. J. Ryan and E. H. Bonfield, "The Fishbein Extended Model and Consumer Behavior," Journal of Consumer Research, 2(September 1975), 118-136.

M. J. Ryan and M. J. Etzel, "The Nature of Salient Outcomes and Referents in the Extended Model," in B. B. Anderson (Ed.), Advances in Consumer Research, Volume III (Cincinnati: The Association for Consumer Research, 1975), 485-490.

R. L. Schmidt, "Implications of a Measurement Problem for Expectancy Theory Research," Organizational Behavior and Human Performance, 10(1973).

M. Sherif, "Group Influences upon the Formation of Norms and Attitudes," in E. E. Macoby, et al. (Eds.), Readings in Social Psychology (3rd ed.) (New York: Holt, Rinehart and Winston), 1958.

A. E. Siegal and S. Siegal, "Reference Groups, Membership Groups, and Attitude Change," Journal of Abnormal and Social Psychology, 55(1957).

L. G. Warner and M. L. DeFleur, "Attitudes as an Interaction Concept: Social Distance as Intervening Variables Between Attitudes and Action," American Sociological Review, 34(1969), 153-169.

A. W. Wicker, "Attitudes Versus Actions: The Relationship of Verbal and Overt Behavioral Responses to Attitude Objects," Journal of Social Issues, 25(1969), 41-78.

THE MARKET FOR PERSONAL GROWTH SERVICES

Philip Kotler, Northwestern University
Lenore Borzak, Northwestern University

Abstract

Americans are enrolling in growing numbers in programs promising them "personal growth." These programs are supplied by (a) "mom-and-pop" enterprises, (b) proprietary enterprises, and (c) growth centers. Four proprietary enterprises--TM, EST, Scientology, and Silva Mind Control--attract and satisfy consumers through meeting a variety of manifest and latent needs. Underlying these services are a hierarchy of self-actualization needs that are becoming more salient in post-industrial society.

Introduction

Personal growth services, broadly defined, includes religious services, psychiatric services, psychological services, self-improvement services, inspirational services, and even occult services. It could include the Moonies, Total Woman, Alcoholics Anonymous, and even Weight Watchers. We will confine our attention to a narrower band of personal growth services that have the following characteristics:

1. The service is bought. This excludes traditional and contemporary religious movements, both Western (e.g., the Unification Church, Campus Crusade for Christ) and Eastern (e.g., Hare Krishna). Members of these religious movements do not see themselves as buying a service so much as living out a faith.
2. The service is available to anyone who wishes to buy it. The personal growth services are available to the mass market. Organizations that offer enlightenment programs only to their own members are excluded.
3. The service consists primarily of group processes rather than of individual counseling. This eliminates individual psychotherapy, religious counseling, and paid advice systems.
4. The seller is considered professionally qualified to offer the service. The founder has professional credentials and/or exemplifies to a high degree the qualities that the service offers to impart.
5. The service is trademarked and marketed. The seller creates an identity for the service and markets it.

Personal growth services are supplied to the market in three forms. A "mom-and-pop" supply system consists of a therapist or team conducting group workshops at regular intervals for a fee. They may be conducted in a home, office, or hotel. If they are not trademarked, they are at least closely identified with a particular supplier. There are hundreds of entrepreneurs running personal growth workshops on this basis.

A proprietary (or "franchised") supply system consists of an entrepreneur designing and copyrighting a particular service, developing licensed trainers to deliver this service in different cities, and taking responsibility for marketing and financing the growth of the enterprise. Scientology, Silva Mind Control, Transcendental Meditation (TM), and est (Erhard Seminar Training) are the best examples.

Finally, a growth center supply system consists of a single or multiple site center that offers a wide variety of courses on a regular basis that may run from one day to over several months. Such centers are located in urban areas (such as Esalen in San Francisco or Oasis in Chicago) and in Arcadian settings (Esalen at Big Sur, California, or Naropa Institute near Boulder, Colorado). Esalen, the earliest and best known growth center, issues a catalog every four months listing over one hundred personal growth courses, bearing such titles as "Self Acceptance--the Freedom to Be," "Psychosynthesis: Choosing to Change," and "Getting Out of Your Trap." Each course is offered either by regular Esalen staff members or well-known therapists. In recent years, Esalen has taken some of its programs on the road offering intensive training weekends in select cities.

How many Americans have taken these personal growth services? As of 1976, Scientology claims 5,000,000 registrants; Silva Mind Control, 435,000; Transcendental Meditation, 250,000; and est, 100,000. (Lande, 1976) Figures are not available for mom-and-pop enterprises and growth centers, but one can safely guess that they have serviced several million American consumers.

Though each supply system could be explored at length, we shall focus on the four proprietary services listed above, since these four have been the most conspicuous in the last few years, primarily in terms of numbers of consumers served. If, as Adam Smith has suggested in Powers of the Mind, TM is the McDonald's of the meditation business, with its relatively low fixed price, standard item, and increasing numbers of franchises or outlets, then Scientology, Silva Mind Control, and est are the McDonald's of their respective segments of the market. For highly mobile Americans who may reside in a particular area on an average of less than three years, these proprietary services insure quality control by selling a uniform item. In est, for example, the same nine trainers conduct the training throughout the United States, and there is virtually no difference between taking the training in Washington, D.C. or in San Francisco. The post-training seminars are also nearly identical in all the est centers.

Data on the characteristics of consumers are not available. However, most observers agree that the consumers are largely middle-class, college educated, and white. Younger persons predominate although all ages are represented. The various services draw both sexes in about an equal ratio. One journalist suggests that such services attract people who have "made it," because once they make it, they realize that they're still not happy, and they go looking for something else. (Preston, 1976)

Another factor contributing to the preponderance of white-collar consumers is that an increasing number of corporations are willing to pay the cost of personal development workshops for their executives. Thus, these proprietary services become part of the expense-account economy.

When did these personal growth services first appear? One could argue that they originated at the very beginning of human society. Early mankind's fears and vulnerabilities led to the need for services to counsel

and console him. Myths were created, religions built, and practitioners such as priests, shamans, and witch-doctors came forth to offer inner peace in return for allegiance or fees. Whether the incentive to seek help was prompted by the desire to exercise an evil spirit or to rid the body of an excess of one of the four "humours" that upset the temperament, people have paid attention to the working of their "inner lives." However, the practitioner in the past was called in to rectify what was wrong with the person. Attention to the inner life, which reached a new high with the emergence of psychiatry in the late nineteenth and early twentieth centuries, typically has emerged from an illness model. Freud and other physicians developed the concepts of modern psychiatry on the basis of new insights into the pathology of the human mind. They listened to patients describe their mental and emotional problems and prescribed a course of therapy to treat their symptoms and effect a cure. The focus was on disorder, on what went wrong in the past that could be corrected in the present.

Paradoxically, on the one hand, we can trace the modern roots of the human growth services to psychiatric medical practice. The patient's talking to a sympathetic and alert physician was developed into a therapeutic concept. One researcher coined the term "therapeutic man" to describe the prevalence of this healing outlook in the twentieth century. (Rieff, 1966; Hopkins, 1961)

On the other hand, the Freudian illness model can be seen as antithetical to the personal growth industry, especially if we turn to humanistic psychotherapy and all that its practitioners spawned as another and perhaps closer source of these growth services. Many of the so-called "humanistic" therapists were trained in the Freudian tradition. Yet, with few exceptions, they were clinical psychologists rather than physicians and were not committed to a medical orientation. Instead of looking at the human psyche to see what was wrong, they looked at what was right and healthy about human beings and what prevented people from realizing or actualizing their full potential.

Both Abraham Maslow and Carl Rogers, two of the most influential of the humanistic psychologists, felt that Freud and his disciples had built up theories of human nature based on the study of men and women at their worst. In the early 1940's Rogers pointed out how his client-centered approach differed from Freud's:

This newer approach differs from the older one in that it has a genuinely different goal. It aims directly toward the greater independence and integration of the individual rather than hoping that such results will accrue if the counselor assists in solving the problem. The individual and not the problem is the focus. The aim is not to solve one particular problem but to assist the individual to grow, so that he can cope ... in a better-integrated fashion ... therapy is not a matter of doing something to the individual ... it is instead a matter of freeing him for normal growth and development, of removing obstacles so that he can again move forward. (Rogers, 1942)

For Rogers "growth" is synonymous with health, and both underlie man's basic nature: the one basic tendency and striving is "to actualize, maintain, and enhance the experiencing organism." (Rogers, 1951)

This is a view of man as basically good rather than evil, a being whose nature is positive. It is a view that is explicitly contrasted with Freud's view that the id, man's basic and unconscious nature, is primar-

ily made up of instincts which would, if permitted expression, result in incest, murder, and other crimes. For Rogers the basic nature of the human being, when functioning freely, is constructive and trustworthy. He quite consciously substituted the term "client" for "patient" to depart from the Freudian illness model.

Abraham Maslow was also influential in creating what we shall henceforth call a wellness model for human services. Drawing upon his own findings both within and outside of the context of the therapeutic situation, Maslow defined what he called a viable third alternative to behavioristic psychology and to orthodox Freudianism. He labelled it a Third Force "health-and-growth psychology" grounded on the assumptions that man is basically good, and that he seeks to actualize his potentialities. Certain restraints or self-imposed barriers may prevent him from reaching his full potential, including energies directed to lower order needs such as providing food and shelter for himself. Once such barriers are removed, the person can be wholly and fully human and is able to make what Maslow calls the "growth choice" to be actualized. (Maslow, 1965, 1954, and 1962)

Both Rogers and Maslow had tremendous impact on a number of psychologists who seized on the notion of the human potential as an untapped natural resource. While there were many thinkers and practitioners from a myriad of sources, eventually this third force thrust came to be labelled the human potential or personal growth movement.

Parallel to this and consonant with the wellness model was the beginning of T-group or sensitivity training in 1946. Kurt Lewin and his co-workers had been working out the theory of using group processes to (1) train leaders; (2) develop self-insight in individuals; (3) create awareness in individuals of how groups function; (4) change attitudes; (5) resolve conflict. The National Training Laboratories (NTL) were set up in Bethel, Maine, and worked principally to train individuals to be more effective as group members, usually in business or industrial settings. The emphasis in the training was on present behavior and an analysis of it in the workshop that could be transferred to back-home settings.

The training group or T-group was the specific activity of NTL and its staff. However, their approach to group-methods became extremely popular and was carried forward into schools and growth centers around the country. Eventually the terms sensitivity training and encounter became synonymous with T-groups, although the links between them were loose. Being sensitive to one's own and others' feelings and encountering and coming to terms with such feelings in the present were practices in T-groups and in the workshops created at places like Esalen Institute. Esalen, established in California in 1962 and offering an ambitious potpourri of experiences in Eastern religion, body awareness growing out of Reichian bioenergetics, and small group processes à la Maslow, Rogers, and the gestalt therapist Fritz Perls, is clearly the grandfather of growth centers.

What is interesting about the rapid proliferation of the centers is that they seem to have been conceived as both a means of improving one's life and a leisure-time activity. Back refers to them as "psych-resorts." (Back, 1972) He notes that they provided a peculiar leisure-time experience accompanied by a feeling of accomplishment.

This latter component is the link to still another source of the personal growth industry and its success in the United States: the "self-improvement" ideology so deeply embedded in American character. Ben Franklin epitomized the American who systematically observed his own actions, evaluated them against certain standards and goals, and took a definite course of action to improve himself. In the nineteenth century, many Americans flocked to various meetings and campgrounds where they would pay a fee to hear lectures by such persons as Ralph Waldo Emerson and Mark Twain on how to think and feel better.

In the twentieth century, millions of Americans would read Dale Carnegie's How to Win Friends and Influence People and Norman Vincent Peale's The Power of Positive Thinking. Tens of thousands would sign up in follow-up courses to master these and other principles of self-improvement.

Closely allied to the concept of self-improvement is another American ideal exalted by Emerson: self-reliance. His prescription for the good life was to be "self-realizing and self-directed," (Emerson, 1946) a notion which is close to David Riesman's "autonomous" person a century later in The Lonely Crowd. While Thoreau may have epitomized the totally autonomous life in Walden, in the twentieth century millions of Americans would live it in a modified version by means of "do-it-yourself" projects.

For the outer life, doing it yourself means learning from others what are the proper tools and ways to proceed, then plunging ahead. In the business of the inner life, the process is the same. The personal growth industry appeals to the "pulling oneself up by the bootstraps" mentality: people pay for a course of training that gives them the tools to live a better life.

In the four proprietary services we are highlighting, the thrust is to give the individual more control over his or her own life. The instruction occurs in a relatively brief period of time, under the tutelage of a trainer-guide who carefully tries to establish rapport as an equal. For example, in est the trainer tells his audience more than once that he is simply a graduate who does the training and is in every way like the trainees. This is very different from the asymmetric doctor/patient relationship in the medical psychotherapeutic model where the individual is ministered to by others.

The wellness model of the humanistic therapies, along with the offshoots that have been lumped together as the human potential movement, can be viewed as consonant with the melioristic streak in American culture. These phenomena all turn on the notion of the perfectibility of human nature. Moreover, several factors seem to make proprietary services such as TM, est, and Scientology acceptable and popular to large numbers of Americans in ways that no therapy has ever achieved.

One is that there is no stigma of mental illness associated with these services. One does not have to be sick to get better. Another is the weight of the numbers of people served. Like McDonald's continually changing marquee announcing how many hamburgers have been served, each of these services takes pride in publicizing how many people have been trained. Perhaps there is a comfort in knowing that Scientology has trained five million people and in imagining that they must be getting something for their money.

Finally, there seems to be safety in numbers. Each of these services has gone far beyond the parameters of

group treatment with fewer than 20 persons that neo-Freudian therapies developed. An est workshop with 250 people in the ballroom of a hotel is really closer in origins to a nineteenth century revival meeting than to group therapy. This is mass distribution to the mass market.

While there may be religious undertones or the elitism associated with the mystique of belonging to the ever-growing throng of graduates or mediators, an important part of the appeal of these services is to a secular and science-believing generation of Americans. At introductory lectures TM distributes reprints of articles from prestigious scientific journals to document the physiological benefits of meditating. Scientology's professional aid, the "E-Meter" is a scientific-looking instrument with a moving needle, and its methodology, dianetics, is presented as a science for achieving total mental health. Silva Mind Control contains elements of alpha brain-wave training, and est draws upon Scientology and physics in its lectures.

This appeal, combined with the intensity of the group experience, is a potent combination for consumers. Perhaps, as Back suggests, such personal growth services meet many needs formerly met by religion, giving an air of scientific respectability to feelings people have previously sought in a religious experience.

Needs and Wants in the Personal Growth Services Market

What are people seeking when they sign up in personal growth programs? To answer this we will refer to the hierarchy of needs postulated by Maslow. (Maslow, 1954) He proposed that the most potent human needs are physiological -- hunger, thirst, the need to survive. When these needs are satisfied, people turn their attention to needs for safety -- saving for the future, building better shelters, practicing ordinary prudence. When their safety needs are under control, people begin to pay more attention to their needs for belongingness and love -- to have a family and friends, to love and be loved. At the next level, people strive for esteem and status -- to gain recognition, applause, and even envy in the eyes of others. Finally, when people are secure in these needs, they are free to deal with a need for self-actualization -- to improve their own capacities for thinking, feeling, and doing, experiencing fully and vividly with full concentration and total absorption.

In highly industrial societies, many citizens have achieved affluence. They have left behind their physiological and safety needs and are struggling with higher-order psychological needs. A large number are involved in searching for self-actualization. Within self-actualization, there is itself a nested hierarchy of needs.

The first need is for relaxation. The hectic pace and pressure in industrial society have strained people's nerves and created a billion dollar pain-killing industry featuring aspirin, valium, and other medicinal "calming" agents. Personal growth services represent an alternative or supplement to pharmaceutical means of detensification. Probably over eighty percent of the population would admit to an interest in learning how to relax more easily. They buy a tremendous number of books on relaxation and life simplification.

Relaxation is perhaps the simplest way to describe this need. A more active form of the need might be called reenergization, where the person wants to overcome tired feelings and regain aliveness of body and mind. Recentring is still a stronger form of the need in which the person wants to feel whole again, to be in

touch with himself or herself. This may take the form of connecting the mind and body in new ways. Rollo May, a widely-read humanistic psychologist, has stressed people's need for centering as a means of personal integration, finding a center of strength within the self particularly in the twentieth century age of anxiety. (May, 1953) TM is the primary service catering to the relaxation class of needs. The other services supply relaxation techniques but not as their primary mission.

The next self-actualization need is for awareness expansion. Many people sense a condition of not being fully aware of their own feelings, either denying, avoiding, or suppressing them. They also sense barriers to "being" with other people in the "here and now," that is, being empathic. Thus awareness expansion takes the form of wishing to enhance one's self-awareness and other-awareness, and such human growth services as est and Scientology are tuned in to this need.

A more active need is for enhanced interpersonal effectiveness. Many people would like more mastery and impact in social situations. They envy those who come on strong and who know how to get what they want. Dale Carnegie training and assertiveness training are among the services offered to this market segment.

The next need is for personality change. Many people are unhappy about their emotional makeup. They cry too much, are angry too much, are too hostile, or have no feelings at all. Some want to achieve more spontaneity; others, deeper feelings; still others, evenness of temper. They are in the market for Gestalt therapy, rational therapy, and primal scream therapy, all promising to recondition their emotional makeup.

A need that is emerging in recent years is a desire for self-transcendence. Instead of investing more in their ego, some people are trying to reduce or renounce their ego as the key to becoming tranquil, compassionate, and satisfied with their lives. They seek to attain higher levels of consciousness and enlightenment. They respond to gurus and join ashrams, camps, and monasteries in their search for self-transcendence.

Finally, some people seek to transcend their ordinary state of consciousness through occultism. They want to believe in levels of reality beyond those obvious to the senses. They are in the market for extrasensory perception, levitation, clairvoyance, communication with the dead, materializations, astral-projection, ghosts, and other occult occurrences.

Entrepreneurial Characteristics of Personal Growth Service Firms

The majority of human growth service enterprises are created by a charismatic individual rather than by a group of individuals whose identities remain vague. This is especially true of the proprietary services. Scientology centers around Ron Hubbard, TM around the Maharishi Mahesh Yogi, Silva Mind Control around Joseph Silva, and est around Werner Erhard. The leader is usually very bright, highly articulate, dedicated, very visionary. He attracts enthusiastic supporters who give their time and energy at little or no pay. The leader formulates the core service package and it goes through additional refinements in the process of being presented to different groups. The current form of the service represents the accumulated experience of the leader and several licensed trainers as to what works and what doesn't work.

To date, most personal growth firms have gone after the middle majority market because of its high discretionary income and high interest in self-actualization. These enterprises have typically sought both sexes, although a few new services are aimed specifically at the women's market (e.g., Total Woman, Applied Potential). Most services are geared to adult consumers, although some firms have designed segmented services for teens or senior citizens. Most services are available to people regardless of occupation, although some firms have designed segmented versions for special groups such as professors, executives, or prisoners. As the market evolves, undoubtedly firms will start targeting their services to mini-markets of certain social classes, races, and occupations.

Several of the proprietary services address themselves to meeting one particular need in the self-actualization hierarchy. TM, for example, positions its service primarily in the relaxation market. Here is a recent statement (from a TM poster on a college bulletin board):

Transcendental Meditation is natural, effective, systematic, refreshing, simple, effortless, spontaneous, easily learned, scientifically verifiable, practiced twenty minutes morning and evening to develop the full potential of the individual. It isn't a lifestyle, self-hypnosis, concentration, contemplation, mind control, a philosophy, a yoga exercise, an intellectual practice, a religion, a diet, a special way of dressing, or difficult to learn.

On the other hand, Silva Mind Control promises the consumer a host of benefits (from a brochure distributed by Silva Mind Control):

Control of your own mind--are you ready for it? Are you ready to do whatever you do, better? Are you ready to be a better person? Student, housewife, gardener, businessman, teacher, parent, doctor, salesman, politician, or even free spirit? Do you want to understand your self and others better? Are you ready to listen to your intuition when it tries to tell you something? Are you ready to remove those limits you keep placing on yourself? Are you ready to reach for your inner strengths and talents and make them work for you? Are you ready to control the way you think about your problems? And solve them better? Are you ready for an incredible trip into your own mind?

Scientology is also quite broad in the range of needs it alleges serving. It offers a succession of courses to help the person advance to higher grades of ability:

- Grade I. A person acquires the ability to recognize source of problems and make them vanish.
- Grade II. A person acquires the ability to achieve relief from the hostilities and suffering of life.
- Grade III. A person acquires freedom from the upsets of the past and ability to face the future.
- Grade IV. A person learns how to move out of fixed conditions and gain ability to do new things.

Grades V, VI, etc.

The personal growth service founder designs a program consisting of a mix of manifest and latent services. Manifest services are the stated benefits that the supplier promises to deliver to the target market. TM offers to teach a method that can produce relaxation and inner peace. Est offers a training that will "transform your ability to experience living so that the situations you have been trying to change or have been putting up with clear up just in the process of life itself."

Latent services are the unstated benefits that participants may seek or expect to find in the program. Most personal growth entrepreneurs are acutely aware of the "hidden agenda" of latent needs that their services might cater to, including:

1. Need to relieve boredom. Many people find their life situation drab and see a personal growth program as a way to spend their time in a more interesting manner.
2. Need for sociability. Many people are seeking an opportunity to be with other people, make new friends, and possibly develop romantic relationships, all in a socially acceptable context.
3. Need to be interesting to others. Many people want to be regarded as having done something interesting and different. They may see the service as a way to become more interesting to other people.
4. Need to find community. Many people do not enjoy their current group affiliations and want to become part of a new family or community with which they could share their energy, time, and love. The human potential group provides opportunities for volunteering services and deepening friendships.
5. Need to find meaning in life. Many people who sign up for these services have abandoned traditional religion and are looking for something new in which to believe.

Because of the multiple motives that personal growth programs satisfy, participants come away with quite different ways of describing what they got. It is common for est graduates, for example, to emphasize quite different benefits and insights that they gained from their training. Each gets "It" differently. One cynic has suggested that it really does not matter what the content of these programs is in most cases. Just putting people together to share an intensified weekend will produce a "high."

The Future of Personal Growth Services

Personal growth service firms are presently operating in a boom market. The growth of this market will attract more suppliers. Growing competition will lead firms into increased market segmentation and specialization. Eventually the rate of market growth will slow down, forcing firms to undertake more intense marketing. This will express itself in the accelerated development of new services, the intensified use of advertising and personal selling, and the cultivation of more publicity. Some suppliers will be forced out of the market. Eventually the market will stabilize at a certain size and with a certain well-known set of leading suppliers.

Some observers expect the whole size of the market to shrink as people get overexposed or lose their interest in ego services. They see the human potential movement as faddish. Our belief, on the contrary, is

that personal growth services meet a permanent need in the marketplace. They will be around as long as no other means, such as new mind drugs or psycho-technical products, are found to be more effective.

References

Kurt Back, Beyond Words (New York: Russell Sage Foundation, 1972), p. 131.

Ralph Waldo Emerson, "The American Scholar," The Portable Emerson (NY: Viking Press, 1946), p. 37.

Nathaniel Lande, Mindstyles/Lifestyles (Los Angeles: Prive/Stern/Sloan, 1976, pp. 128, 135, 147, 298.

Abraham Maslow, "Self Actualization and Beyond," in J. Bugental (ed.), Challenges of Humanistic Psychology (NY: McGraw-Hill, 1965).

Abraham Maslow, Motivation and Personality (NY: Harper & Row, 1954).

Abraham Maslow, Toward a Psychology of Being (NY: Van Nostrand Reinhold, 1962).

Rollo May, Man's Search for Himself (NY: WW Norton, 1953).

Marilyn Preston, "The Elusive Search for Inner Self," Chicago Tribune, September 21, 1976.

P. Rieff, The Triumph of the Therapeutic (NY: Harper, 1966).

Carl Rogers, Counseling and Psychotherapy (Cambridge, Mass.: Houghton-Mifflin, 1942), pp. 28-29.

Carl Rogers, Client-Centered Therapy (Boston: Houghton-Mifflin, 1951), p. 487.

SPARKS AND TUCKER REVISITED: A REANALYSIS
AND REPLICATION

Charles M. Schaninger, University of Massachusetts
V. Parker Lessig, University of Kansas

Abstract

The personality and product usage measures employed in the Sparks and Tucker (1971) study were replicated on a sample of 142 male college students at a large midwestern university. A comparison of the simple correlation matrices between the two sets of data revealed only limited similarity. Both sets of data revealed canonical roots of similar magnitude and significance, although only limited similarity of canonical loadings among the canonical functions was demonstrated.

Introduction

Within the field of personality and product usage, most of the research appearing since the late 1960's has been strongly influenced by Sparks and Tucker's (1971) study. Prior to this and Kernan's (1968) study, most of the work in this area had focused on either simple univariate analyses, or on single criterion, multiple predictor studies. Canonical correlation analyses, cluster analyses, and multiple discriminant analyses have been used in a large number of studies (Alpert, 1972; Greeno, Sommers, and Kernan, 1973; Darden and Reynolds, 1974; Villani, 1975). Essentially, these studies suggest the existence of "molar personality types" -- i.e., patterns of product usage across product classes associated with corresponding personality profiles. Personality is viewed in a gestalt sense, with various combined multivariate approaches used to identify the different molar types. Many studies have utilized similar multivariate-statistical approaches with AIO/lifestyle/psychographic data, to identify "molar types" or segments in the market. Books by King and Tigert (1971), and Wells (1974), as well as papers by Wells (1975) and Kassarjian and Sheffett (1975), describe this vein of research.

The Sparks and Tucker paper had a strong impact because it took a broadened conceptual approach toward the area of personality and product usage, and because it's conclusions suggested the existence of much stronger and more meaningful relationships than had been found in previous investigations.

A number of authors (Kollatt, Blackwell and Engel, 1972; Engel, Kollatt and Blackwell, 1973; and Robertson and Ward, 1973) have stressed the need for replication of consumer behavior studies, particularly for studies which have a significant impact and which used limited samples. Our objective was to replicate (using new data) and to reexamine (using Sparks and Tucker's original data) the Sparks and Tucker study. We were specifically interested in examining the degree to which the original and the replicated studies yield comparable results.

Methods

The subjects for the replication were 155 male undergraduate college students enrolled in introductory marketing courses at a large midwestern university. Data were collected for the same eight personality variables and seventeen product-usage variables used by

Sparks and Tucker. Of the 155 subjects comprising our total sample, 142 accurately completed all questions. This sample is roughly comparable to that of Sparks and Tucker's original study -- although the samples do come from universities located in different geographical areas and from different time periods.

We obtained and reanalyzed the raw data for the 173 subjects completing all questions for the original Sparks and Tucker study.¹ We found a number of minor discrepancies between the correlations based upon reanalyzing their data and those reported in the original article. Evidently their calculations were based on pairwise deletions of missing data on their entire sample of 190 subjects, rather than on the reduced sample of fully completed forms for the 173 subjects serving as our reanalysis sample.

The relationships between personality and product usage were examined using the two approaches presented by Sparks and Tucker: an examination of the simple correlations between each personality trait and product usage score, and an examination of the canonical correlation analysis which focused on the overall association between the two sets of variables (product usage and personality). For both the original and replication samples, the overall strength of association between the two sets of variables was examined by Bartlett's test of canonical correlation significance, by Veldman's test of canonical correlation significance, and by Stuart and Love's (1968) redundancy measure.

Finally, the nature of the relationships found between the personality and product usage measures for the two samples was compared by examining the factor loadings associated with each set of linear canonical functions.

Results

Simple Correlations

From an examination of their correlation matrix, Sparks and Tucker reported "some weak and spotty relationships between personality traits and particular product use." They found 18 of the 136 cross correlations to be significant at the 0.05 level. Our reanalysis of Sparks and Tucker's data (n = 173) also produced 18 significant correlations; 13 of these were the same as Sparks and Tucker found. Only 12 correlations were significant on the replication sample; of these only five were the same as those of the original study. These five were between the following pairs: ascendancy and use of alcoholic beverages, ascendancy and fashion adoption, sociability and use of alcoholic beverages, sociability and fashion adoption, and between cautiousness and use of alcoholic beverages.

¹These data were made available through the courtesy of Professor Sparks.

Multivariate Relationships

Table 1 presents the first three canonical roots for our reanalysis of Sparks and Tucker's data (n = 173), for the replication study (n = 142), and those reported in Sparks and Tucker's article. We note that in both the original Sparks and Tucker study and the reanalysis of their data by Lambert and Durand (1975), it was concluded that the first three roots are significant. However, the third root in both studies was significant at the 0.01 level and not at the 0.05 level. Based upon their observations, Sparks and Tucker interpreted the first three sets of linear canonical functions. It should be noted that the studies by Sparks and Tucker and by Lambert and Durand applied the significance test developed by Veldman (1967), rather than Bartlett's (1941) test. When Veldman's test was applied to Sparks and Tucker's data (n = 173, not 190) and to our replication data, only the first two roots were significant. Examination of the Bartlett's test revealed only the first canonical root to be significant for each set of data. (These chi-square values and their corresponding degrees of freedom are also reported in Table 1.)

The discrepancy between the Veldman and the Bartlett procedures concerning the number of significant canonical roots should be noted. The correspondence between these two tests has been presented by Alpert, Peterson, and Martin (1975). Essentially, Veldman's test

partitions the first Bartlett test chi-square value (of the significance of the first and remaining sets of canonical functions) to calculate the relative chi-square contribution attributable to each of the roots. Disregarding the significance testing procedure employed, both sets of data were comparable with regard to the number of significant canonical roots.

Although both Bartlett's test and Veldman's test provide tests of the statistical significance of the association between two sets of variables, little additional information is provided. Stewart and Love's redundancy test was therefore applied to measure the amount of variation in the set of personality scores. Table 1 also shows the redundancy scores obtained on the Sparks and Tucker data and on the replication data. These scores indicate that 7.37 percent of the variation of product usage is associated with variation in personality for the Sparks and Tucker data, and 8.28 percent for the replication data. The reader is referred to Lambert and Durand (1975) and to Alpert, Peterson, and Martin (1975) for a discussion of this multivariate squared multiple correlation measure. It should be noted that the more complex mathematical form for its calculation simplifies to the mathematical average of the multiple R² for each criterion variable with the set of predictors.

TABLE 1
First Three Canonical Roots

Sparks and Tucker Reanalysis (n = 173)				
Canonical Root	Chi-Square (Bartlett's)	d.f. (Bartlett's)	Chi-Square (Veldman's)	d.f. (Veldman's)
0.348	188.059 ^b	136	68.480 ^b	24
0.293	119.579	112	55.456 ^b	22
0.119	64.123	90	20.347	20
Redundancy score, $R^2_{c/p} = 7.37\%$				
Replication Data (n = 142)				
Canonical Root	Chi-Square (Bartlett's)	d.f. (Bartlett's)	Chi-Square (Veldman's)	d.f. (Veldman's)
0.330	171.725 ^a	136	51.335 ^b	24
0.258	120.390	112	38.184 ^a	22
0.179	82.206	90	25.349	20
Redundancy score, $R^2_{c/p} = 8.29\%$				
Sparks and Tucker's Original Study (n = 190)				
Canonical Root			Chi-Square (Veldman's)	d.f. (Veldman's)
0.367			72.742 ^b	24
0.300			56.703 ^b	22
0.171			29.842	20

^aSignificant at the 0.05 level.

^bSignificant at the 0.01 level.

Nature of the Relationships

Table 2 presents the factor loadings for the first and the second set of canonical functions for our reanalysis of Sparks and Tucker's data and for our replication sample. It should be noted that loadings represent correlations between the original variables and the canonical variate scores, and not the canonical weights (coefficients) used to multiply standardized variable scores calculating canonical variate scores. In several instances, both the direction and magnitude of the weights were different from the loadings for the reanalysis and for the replication data. When collinearity existed, the weights tended to be lower than the loadings in magnitude and when a given variable had a suppressor effect, its weight was opposite in direction to the loading.

An examination of the first set of canonical loadings for the two samples demonstrated that the contributions made by the original variables in explaining the first set of canonical relationships differ. An interpretation of these loadings for the Sparks and Tucker data suggests a pattern of heavy usage of cologne and alcohol, and high fashion adoption associated with high sociability and ascendancy, and with low responsibility and cautiousness. The comparable first set of canonical functions for the replication sample suggested a pattern of low shampoo usage and high coffee usage associated with higher scores on original thinking, cautiousness, and personal relations. Both conclusions

may be viewed as intuitively appealing, but neither could be described as representing the same basic underlying relationship.

The second set of linear functions from the Sparks and Tucker data demonstrated a product usage pattern consisting of heavy usage of mouthwash, headache remedies, and coffee associated with lower emotional stability, vigor, and original thinking, and higher cautiousness scores. This interpretation, based on loadings, is basically consistent with that given by Sparks and Tucker, based upon their examination of their canonical coefficients (weights). It should be noted that the test of emotional stability, from the Gordon Personal Profile, tends to load highly on items related to nervousness, rather than to those which would indicate emotional instability in a clinical sense.

The pattern of loadings for the second set of canonical functions for the replication data was also dissimilar to the second function for the Sparks and Tucker data, although relatively similar to the first function for their data. A product usage pattern of heavier consumption of alcoholic beverages, chewing gum, and fashion adoption was associated with lower cautiousness and higher sociability and ascendancy. Although there are differences between the second canonical function in the replication data and the first for the Sparks and Tucker data, a similar underlying pattern tends to emerge.

TABLE 2
A Comparison of the Variable Loadings of the First Two Canonical Functions
for the Sparks and Tucker and the Replication Data

Variables	First Canonical Function		Second Canonical Function	
	Sparks & Tucker Data (n = 173) Loadings	Replication Data (n = 142) Loadings	Sparks & Tucker Data (n = 173) Loadings	Replication Data (n = 142) Loadings
<u>Product Usage</u>				
Headache Remedy	.269	-.015	.440	-.285
Mouthwash	.240	.093	.507	.014
Men's Cologne	.473	.128	.012	.122
Hair Spray	.167	-.284	.127	.087
Shampoo	.324	-.509	-.112	-.273
Antacid Remedy	.191	-.031	.282	-.187
Playboy	.350	-.123	-.314	.285
Alcoholic Beverages	.481	-.192	-.288	.671
Brush Teeth	-.074	-.052	.112	.108
Fashion Adoption	.613	-.270	.035	.374
Complexion Aids	.181	-.244	.055	.048
Vitamin Capsules	.330	.198	.058	.005
Haircut	-.144	.056	-.241	-.214
Cigarettes	.308	.057	-.051	-.097
Coffee	-.012	.599	.397	.057
Chewing Gum	.300	-.265	.130	.331
After-Shave Lotion	.067	.164	-.113	.120
<u>Personality Traits</u>				
Ascendancy	.688	.112	-.129	.610
Responsibility	-.520	.222	-.292	.006
Emotional Stability	-.295	.320	-.602	-.037
Sociability	.844	-.169	.059	.663
Cautiousness	-.433	.446	.368	-.688
Original Thinking	.287	.839	-.373	.273
Personal Relations	-.016	.403	.027	.122
Vigor	.049	.023	-.400	.093

Although the roots were not statistically significant, there was also some degree of similarity between the third root of the replication data and Sparks and Tucker's second root; however, the similarities were not as clearcut as those between the second and first (respectively) roots. Cross loadings were also calculated (the correlation between the individual predictor variables and the criterion canonical variate scores and vice versa). These were, as expected, generally lower, yet basically comparable to the original loadings for both data sets. All of the variables mentioned demonstrated significant loadings (correlations) at the .05 level, and the great majority of the cross loadings were also significant at this level. Although the significance tests of the loadings and cross loadings cannot be interpreted as significant correlations between independent variables, their relative significance levels merit examination in lieu of any other significance measures.

Summary

A comparison of the findings obtained from the reanalysis of Sparks and Tucker's data and the same analysis on the replication data reveals several similarities worth noting. First, as shown in Table 1, the magnitude of the first three canonical roots associated with the two sets of data is very similar. Second, for both sets of data, according to Bartlett's test only the first set of canonical functions is statistically significant. However, using Veldman's approach, the first two sets of canonical functions are shown to be significant for both sets of data. Finally, strong similarities are noted in the redundancy scores obtained between product usage measures and personality traits. The redundancies measured from the reanalysis of Sparks and Tucker's data and from the replication data were 7.37 and 8.28, respectively. These figures compare favorably to the redundancy of 6.2 found by Lambert and Durand.

A comparison of the matrices of simple correlations between product usage scores and personality measures for the two sets of data revealed a similar number of significant correlations, although less than half of these significant correlations were for the same pairs of variables. Direct similarities were not noted between the two sets of data when the factor loadings for the first two canonical functions were compared, although the second canonical function for the replication data revealed a pattern similar to that demonstrated in the first function of the Sparks and Tucker data.

The differences in the simple correlations between those reported in Sparks and Tucker's article, based on the full sample of 190 and those found by reanalyzing the data for the 173 with complete data, as well as the evidence of instability of canonical weights found by Lambert and Durand, suggest that the differences in the findings are due more to violation of the statistical assumptions underlying canonical analysis and to random error in the relatively weak relationships found for these small sample steps than to differences between the two samples.

Based upon the consistency of results across a number of different studies, it does appear that a statistically significant relationship exists between product usage and personality traits when these measures are treated as a whole, i.e., multivariately. Given the magnitude of the redundancy indices, ranging from 6.2 to 8.28 percent, it cannot be concluded that this relationship is significant in a practical sense.

Although the relationships demonstrated by Sparks and Tucker's first canonical function do appear to be supported by the loadings of the second function of the replication data, there were still many inconsistencies between the loadings and percentages across the different canonical functions for the two sets of data. Generalizations as to the specific nature of relationships between personality traits and product usage are questionable, and probably should not be made without stronger evidence than that which has been obtained from an unreplicated study.

References

- Mark I. Alpert, "Personality and the Determinants of Product Choice," Journal of Marketing Research, 9 (February 1972), 89-92.
- Mark I. Alpert, Robert A. Peterson and Warren S. Martin, "Testing the Significance of Canonical Correlations," Combined Proceedings, American Marketing Association, 1975, 117-19.
- M. S. Bartlett, "The Statistical Significance of Canonical Correlations," Biometrika, 32 (January 1941), 29-38.
- William R. Darden and Fred D. Reynolds, "Backward Profiling of Male Innovators," Journal of Marketing Research, 11 (February 1974), 79-85.
- James F. Engell, David T. Kollatt and Robert D. Blackwell, Consumer Behavior (2nd edition), New York: Holt, Rinehart & Winston, 1973.
- Daniel W. Greeno, Montrose S. Sommers and Jerome B. Kernan, "Personality and Implicit Behavior Patterns," Journal of Marketing Research, 10 (February 1973), 63-69.
- Harold H. Kassarian, "Personality and Consumer Behavior: A Review," Journal of Marketing Research, 8 (November 1971), 409-418.
- Harold H. Kassarian and Mary J. Sheffett, "Personality and Consumer Behavior: One More Time," Combined Proceedings, American Marketing Association, 1975, 197-201.
- Jerome Kernan, "Choice Criteria, Decision Behavior, and Personality," Journal of Marketing Research, 5 (May 1968), 155-64.
- Charles W. King and Douglas J. Tigert (eds.), Attitude Research Reaches New Heights, Chicago: American Marketing Association, 1971.
- David T. Kollatt, Robert D. Blackwell and James F. Engel, "The Current Status of Consumer Behavior Research's Development during the 1968-72 Period," Association for Consumer Research Proceedings, 1972, 576-85.
- Zarrel V. Lambert and Richard M. Durand, "Some Precautions in Using Canonical Analysis," Journal of Marketing Research, 12 (November 1975), 468-75.
- Thomas S. Robertson and Scott Ward, "Consumer Behavior Research: Promise and Prospects," in Consumer Behavior: Theoretical Sources, T. S. Robertson and S. Ward (eds.), Englewood Cliffs, New Jersey: Prentice-Hall, 1973, 3-42.

David L. Sparks and W. T. Tucker, "A Multivariate Analysis of Personality and Product Use," Journal of Marketing Research, 8 (February 1971), 67-70.

Douglas Stewart and William Love, "A General Canonical Correlation Index," Psychological Bulletin, 70 (September 1968), 160-3.

Kathryn E. Villani, "Personality/Life Style and Television Viewing Behavior," Journal of Marketing Research, 12 (November 1975), 432-39.

William D. Wells, "Psychographics: A Critical Review," Journal of Marketing Research, 12 (May 1975), 196-213.

William D. Wells (ed.), Life Style and Psychographics, Chicago: American Marketing Association, 1974.

PERCEPTION OF MARITAL ROLES IN DECISION PROCESSES:
REPLICATION AND EXTENSION

E. H. Bonfield, Temple University

Abstract

A study by Davis and Rigaux (1974) was replicated with two changes: greater specificity in describing decision items and use of a nonconvenience, random (cluster) sample. The earlier research was extended by checking for differences between traditional couples (masculine husband with feminine wife) and all couples. With some exceptions, generally attributable to greater item specificity, the Davis and Rigaux results were supported. Differences between traditional and all couples were found.

Introduction

In a recent study, Davis and Rigaux (1974) explored the self-reported influence exerted at three general phases of the consumer decision process. The general theoretical and methodological notions were useful, but two, recognized, shortcomings reduced the overall value of the study. First, the researchers utilized a convenience sample, making questionable any generalizations. Secondly, the 25 economic decisions were, for the most part, general in nature, leading to a possible masking of role specialization at a more "micro" level. For example, Davis and Rigaux defined housing as a single item with "location, purchase price, or rent" in parentheses. However, previous research (e.g., Hempel, 1972) has shown husband dominance, wife dominance, and joint decisions for various aspects of the housing decision. The present replication attempts to correct for these two potential contaminants, thus adding generalizability to the results.

A related question deals with sex roles. In an environment increasingly subject to changes in the perception of sex roles, it is possible traditional couples—masculine husband with feminine wife—have different perceived marital roles in the decision process than couples, for example, in which both mates are androgynous in sex role orientation.

This study was thus addressed to the same two questions as was the research by Davis and Rigaux (p. 51):

1. Do marital roles in consumer decision-making differ by phase of the process?
2. To what extent do husbands and wives agree in their perception of roles at various phases of the decision process?

In addition, this study was also addressed to the question:

3. Do marital roles and husband-wife perception agreements differ among traditional (masculine husband/feminine wife) households as compared to the more general population.

Methodology

Consistent with the replication-extension nature of this study, the methodology closely follows that of Davis and Rigaux (1974).

The Sample

Separately administered questionnaires were successfully completed by both spouses in 60 households selected by means of a cluster sample (Sudman, 1976). Clusters of respondent households were drawn from census blocks in a southeastern SMSA, the blocks and households having been drawn randomly. Only high income, high home value blocks were included in the population. Thus, while the generalizability of the following results is limited to relatively upper social class couples in a single region, the population is specifiable. The relatively high class status of the sample was reflected in the reported occupations of husbands, 61% professional and 25% white collar. While 65% of the wives reported they were homemakers, 16% reported they were professionals while an equal proportion reported they were white collar workers. Nearly half the sample reported annual family incomes of \$30,000 or more. Almost 75% of the husbands and 47% of the wives had earned at least one college degree and an additional 40% of the wives had some college education. The median age of couples of approximately 34 years was apparently greater than in the Davis and Rigaux study since only 12% (compared to 36%) were in their twenties. This age difference was reflected in the family sizes in that fewer had no children, 14% to 25%, and more had at least three children, 33% to 28%.

The Questionnaire

As in the Davis and Rigaux study, each respondent gave information about who was the first to recognize a need, who collected the information about possible choices, and who made the final decision to buy, for 20 decisions in this case. The answer categories in each of the resulting 60 questions were Husband, Wife, and Joint; latter scored 1, 3, and 2 respectively. While the decision items were in most cases the same or similar to those selected by Davis and Rigaux, an effort was made to describe each item as specifically as possible without making the task too tedious for the respondents. Thus, while Davis and Rigaux asked about living-room furniture, for example, respondents in the present study were asked about a sofa for the living room or family room. The complete list of items is included in Table 2, and a comparison of item names between the two studies is included in the Appendix. Use of more specific items limits the generalizability of the results to those items, but increases the degree of generalizability about those items.

The BSRI

Bem (1974) developed the Bem Sex Role Inventory (BSRI) in such a way as to preclude "an inverse relationship between masculinity and femininity (p. 155)." The BSRI contains 20-item masculinity and femininity scales "selected on the basis of sex typed social desirability (Bem, 1975, p. 635)." With the BSRI, then, individuals can be classified as masculine, feminine, or androgynous. Androgyny is the preferred state since, according to Bem, an individual's behavior is limited if he or she can only perform behavior conforming to his or her perception of the masculine or feminine sex role.

While it was an original hope for this study to compare households with androgynous spouses with households with masculine husbands and feminine wives, the number of

total androgynous households was too small (3) for such analysis. Surprisingly perhaps, less than half the households (28) were "traditional" in having masculine husbands and feminine wives.

Findings

These analyses follow the same pattern as Davis and Rigaux (1974), but differences in results as well as traditional-total population differences are noted.

A Framework for Classifying Decisions

Davis and Rigaux utilized a diagram, originally suggested by Wolf (1959), which displays all decisions simultaneously. The vertical axis represents the scale of relative influence between husband and wife and results from the average of scores. The horizontal axis represents the degree of role specialization operationalized in terms of proportion of respondents reporting the decision was joint. Since the axes are not independent, the area of feasible coordinates is limited by a triangle (see Figure 1). While the definitions of the regions of the triangle as husband dominant and wife dominant are self explanatory--although wife specialization and husband specialization might be more appropriate terms--syncratic and autonomic decisions require further explanation. Most simply, syncratic decisions are those in which more than half the respondents reported the decision was joint, while, in the autonomic case, less than half the respondents indicated the decision was joint, but there was not consensus among respondents as to whether the decision was husband dominant or wife dominant. Thus, autonomic decisions, as well as husband and wife dominant decisions, represent role specialization.

Differences in Roles by Decision Phase

Table 1 includes a classification of the 20 decisions at each of the three phases in the decision process. Phase

TABLE I
Patterns of Influence at Three
Stages of the Decision Process
(Percentages)

Pattern of Influence	Phases								
	Problem Recognition			Search for Information			Final Decision		
	D/R ^a	60 ^b	28 ^c	D/R	60	28	D/R	60	28
Husband dominant	8	25	25	12	25	25	8	20	20
Autonomic	40	8	5	36	15	15	20	15	20
Syncratic	28	15	15	24	10	10	52	30	25
Wife dominant	24	52	55	28	50	50	20	35	35

^aDavis and Rigaux (1974) shown in percentages, 73 couples and 25 decisions.

^b60 couples selected by means of a cluster sample.

^c28 of the 60 couples classified as "traditional," that is, masculine husband and feminine wife.

1 is problem recognition, Phase 2 is search and evaluation, and Phase 3 is final decision. Percentages are shown so the results may be compared with those of Davis and Rigaux. The earlier results tend to be supported in that there was little difference between problem recognition and search for information, in terms of proportion in each influence pattern, and there was a marked increase in proportion of syncratic final decisions. However, these differences were not so pronounced as in the Davis and Rigaux study, perhaps because of much higher wife dominance throughout the decision process.

A large proportion of the decisions remained in the same influence pattern through all three phases of the decision process (70% in the present study, 64% in the Davis and Rigaux study). There were no differences between traditional couples and all couples among same pattern processes.

Similar to the Davis and Rigaux findings, husband dominance in the three phases was found for insurance on the husband's life (life insurance in Davis and Rigaux), homeowner's or renter's insurance (other insurance), and lawnmower (garden tools) as well as replacement tires for the primary family car which was not included by Davis and Rigaux. Seven decisions remained wife dominant in all stages, including five which support the Davis and Rigaux findings. The supportive, wife dominant decision patterns included replacement or additional pots and pans for the kitchen (kitchenware), beef roast (food and nonalcoholic beverages), slacks for the wife (wife's clothes), and household cleaning products (described the same in both studies). Also, tooth paste and adhesive bandages decisions were found to be wife dominant while the patterns across the three decision phases for the corresponding, more general items, cosmetics and toiletries and nonprescription drugs and first-aid items in the Davis and Rigaux study, were autonomic-wife dominant-autonomic and all autonomic respectively. Only neck tie for the husband was found to have an autonomic pattern over all three phases in the present study. While it might be argued the corresponding, husband's clothes pattern found by Davis and Rigaux was similar in that the third (decision) phase was only barely syncratic, at no stage, and for neither traditional nor total couples, did the percentage of joint decisions reach 15 for neck ties in the present study, while the percentage of joint decisions may have been as low as 30 only in the first (recognition) phase for husband's clothes in the earlier study. The syncratic pattern characterized decisions about movies (concerts, movies, theater) and family vacation (same) in all phases, thus supporting the earlier findings.

The remaining 6 decisions, which changed patterns across phases, were individually at least qualitatively different. In two cases, Davis and Rigaux results were supported. That is, the patterns for living or family room (other furnishings--rugs, drapes) had wife dominant recognition and search phases and autonomic decision phase, and the pattern for TV for living or family room (TV, Hi-Fi, tape recorder) had syncratic-autonomic-syncratic phases in both studies. Washing machine (household appliances excluding TV) decision processes were similar in both studies being autonomic and syncratic in phases 2 and 3 in both studies, but washing machine need recognition appeared to be wife dominant in the present sample compared to autonomic for the more general appliance category of the earlier study. The remaining 3 cases were quite different. While living room furniture and child(ren)'s toys for birthdays and holidays decision processes were characterized as syncratic at all phases in the Davis and Rigaux study, sofa for living room or family room and children's toys for birthdays were found to be wife dominant in both recognition and search in the present study. Further, in the present study, the sofa decision phase was found to be syncratic

while the children's toys decision phase was found to be autonomic. Finally, car decisions, which went from autonomic to husband dominant to syncratic in the earlier study, were husband dominant, as primary family car, in the recognition as well as search phases in the present study. For the total present sample, the decision phase was syncratic for primary family car, corresponding to the earlier findings, but among traditional couples the decision phase was autonomic.

Davis and Rigaux pointed out, correctly, (p.55):

While an analysis based upon the pattern of influence gives a general impression of the change in marital roles throughout the decision process, the criterion is such that much information is lost. A decision's position could change rather dramatically along the dimension of relative influence as well as in terms of joint decision making and still remain within the same pattern of influence. A much smaller change along either dimension, on the other hand, can produce a shift in patterns if the decision happens to lie close to the boundaries.

The authors next discussed changes in relative influence and joint decisions using averages over all item decisions for each phase. The averaging may have been misleading, however, because the averages were likely to have been biased by the sample of decision items selected for the study. Davis and Rigaux did effectively utilize the figures described earlier to graphically summarize changes between phases 1 and 2 and between phases 2 and 3. Figures 1 and 2 provide the same information for the present data on the total sample and Figures 3 and 4 provide this information for the 28 traditional couples.

The Davis and Rigaux findings of right-facing arrows, indicating greater specialization during phase 2, search for information, was not strongly supported since only half the cases in the present study had arrows pointing in the direction of greater role specialization (see Figure 1) compared to 80% in the earlier study. In two cases, replacement or additional pots and pans for the kitchen (kitchenware) and adhesive bandages (nonprescription drugs and first-aid items), nonspecialization movement between phases 1 and 2 found by Davis and Rigaux was supported. Interestingly, the present study supported nonspecialization for neck tie for the husband (husband's clothes), but the direction was toward specialization among traditional couples. While the direction with respect to children's shoes (children's clothes) were opposite in the two studies, the search phase was strongly wife dominant in both, which does support specialization.

In 5 cases, traditional couples differed from the total sample in terms of direction of the arrows in Figures 1 and 3. Traditional couples moved toward nonspecialization in search for drapes for the living room or family room, tooth paste, and insurance on the husband's life, while the total sample moved toward specialization. The opposite was found for beef roast and movies.

The Davis and Rigaux finding of movement toward joint final decisions--left-facing arrows--was supported as shown in Figures 2 and 4.

Similarity of Roles Perceived by Husbands and Wives

In discussing the analytical advantages of the fact their data were collected from both spouses within the same family, Davis and Rigaux pointed out (p. 55):

SEE TABLE 2
FOR ITEM KEY

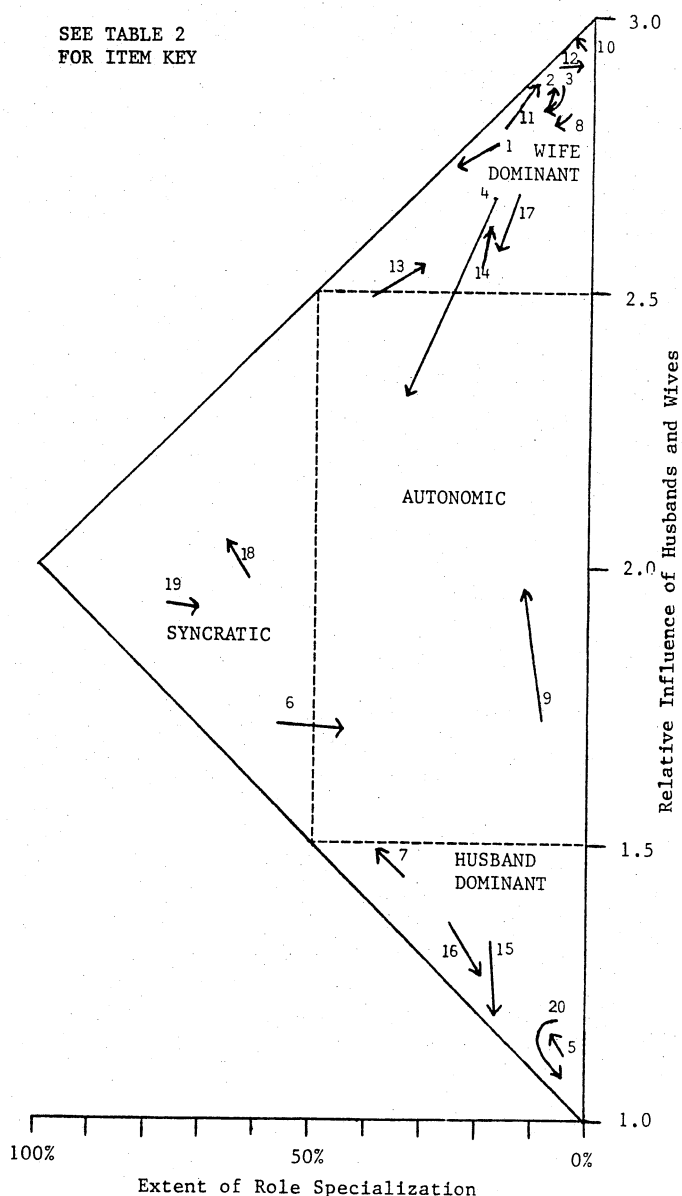


FIGURE 1
Changes in Marital Roles Between Phase 1 and Phase 2
60 Couples (N=120)

. . . it will lead to insights about the more practical issue of which spouse to interview in family research studies . . . Previous research has shown few consistent differences between responses of husbands as a group compared with wives (Wilkening and Morrison, 1963; Davis, 1970; Granbois and Willett, 1970). On the other hand, when intrafamily comparisons of direct questions are made, considerable disagreement is often evident (Scanzoni, 1965; Davis, 1970). While the level of agreement is generally greater than chance, a common finding is that from 10 percent to 50 percent of couples disagree about the influence of one spouse relative to the other for any given decision. The size of this percentage seems to vary with the specificity of questions and, not surprisingly, with the number of response categories.

SEE TABLE 2
FOR ITEM KEY

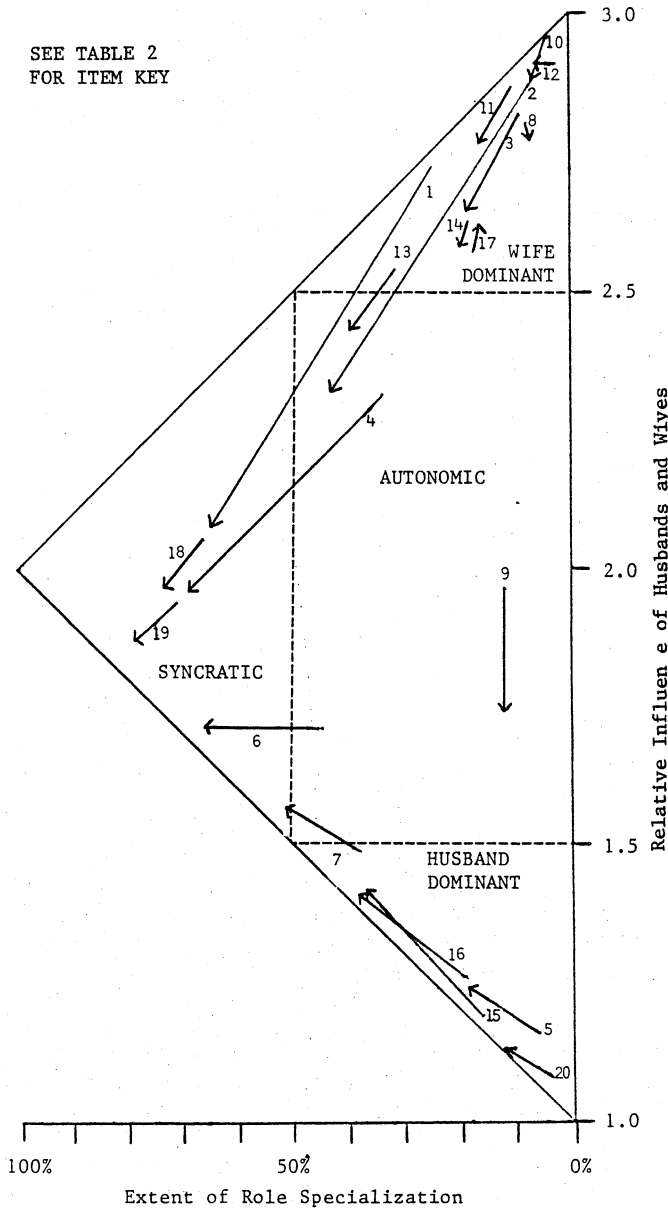


FIGURE 2

Changes in Marital Roles Between Phase 2 and Phase 3
60 Couples (N=120)

in which husbands perceived recognition to be syncratic while wives perceived recognition to be autonomic; and in the cases of phase 2 and phase 3 for primary family car, husbands perceived both search and final decision to be husband dominant while wives perceived these activities to be autonomic. Similarly, in only 4 cases among the traditional couples was there not aggregate agreement between husbands and wives. Among traditionalists, husbands perceived syncratic recognition of need for TV for living or family room while wives perceived recognition as autonomic (syncratic agreement among all couples), husbands perceived autonomic recognition for children's toys for birthdays and holidays while wives perceived wife dominance (wife dominance agreement among all couples), husbands perceived wife dominant search for information on adhesive bandages while wives viewed this activity as autonomic (autonomic agreement among all couples), and husbands viewed the decision about children's toys as autonomic while wives viewed the final decision about children's toys as auto-

SEE TABLE 2
FOR ITEM KEY

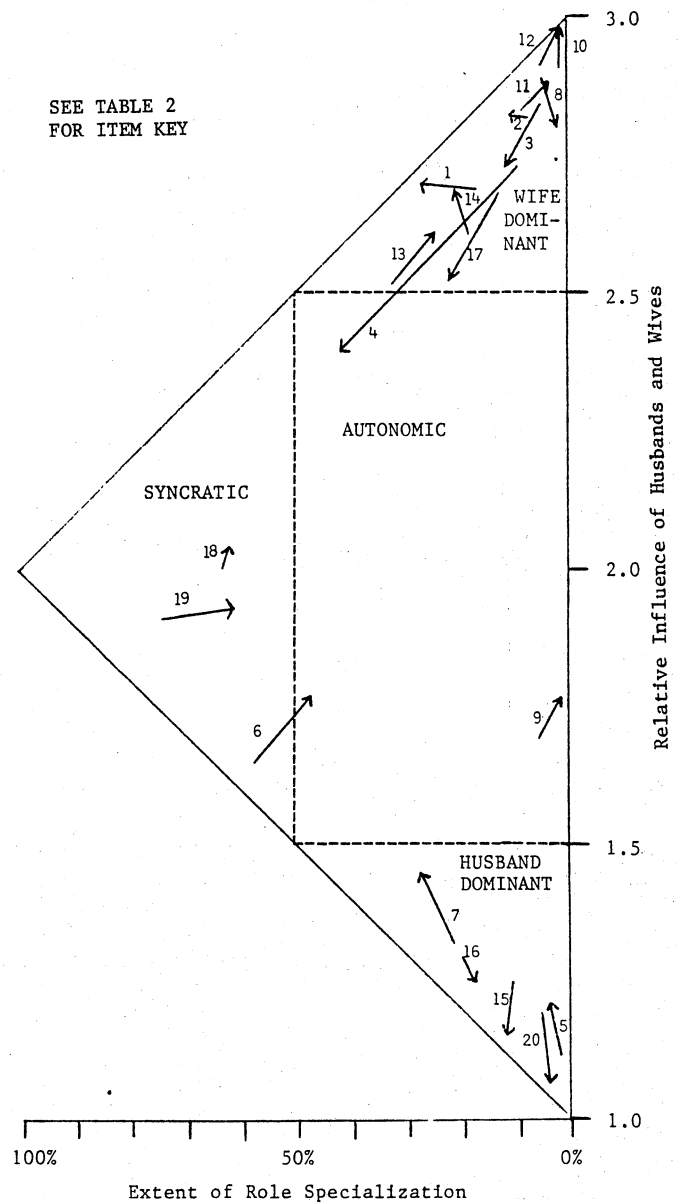


FIGURE 3

Changes in Marital Roles Between Phase 1 and Phase 2
28 Traditional Couples (N=56)

nomie while wives viewed the final decision as wife dominant (autonomic agreement among all couples).

Table 2 provides information about intrafamily agreement in role perceptions. While inspecting average consensus, Davis and Rigaux said their results supported the previous research because, in averaging each decision item across the 3 decision process phases, only 3 of the 25 decisions were classified according to different influence patterns when comparing aggregate husband and wife responses. Without averaging over the 3 phases, similar results were found in the present study. In only 3 of the 60 cases in the present study was there not aggregate husband-wife agreement. Nonagreement was found in the case of family vacation in the first phase or agreement, among couples would not normally be meaningful since such summary statistics may be biased by the selection of decision items, averages were useful in this case because they provided a means of obtaining implications about item specifications and sample types.

SEE TABLE 2
FOR ITEM KEY

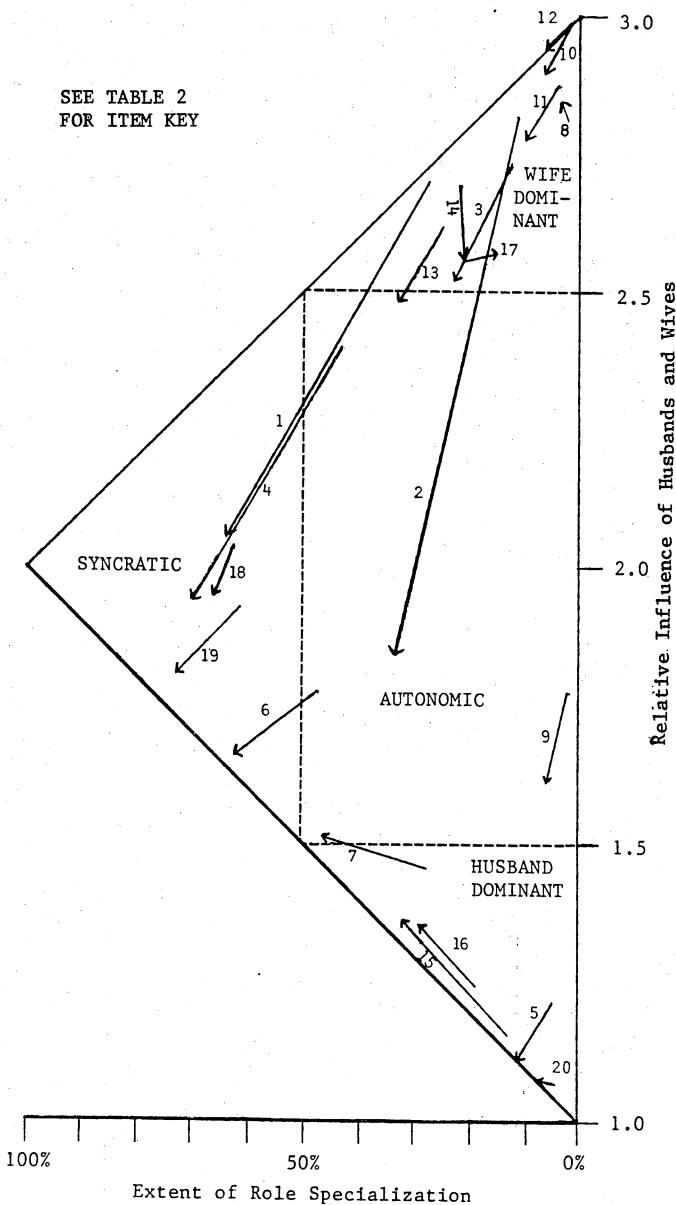


FIGURE 4

Changes in Marital Roles Between Phase 2 and Phase 3
28 Traditional Couples (N=56)

The results paralleled those of Davis and Rigaux in that the percentage of all couples who agreed about their perceived role--consensus--was relatively stable from phase to phase (72, 73, and 68%). However, the average proportion of agreement was slightly, but consistently, higher in each phase than found by Davis and Rigaux (69, 69, and 66%), thus suggesting very slightly greater consensus for more specifically defined consumer decision items. In the aggregate sense, no real difference existed between the 28 traditional couples and the total sample of 60 couples, although the relatively sharper decline in average consensus found at the final decision phase (65% for traditionalists) may be noteworthy. In general, the previous finding that high consensus decisions were those in which relative influence was very skewed (e.g., children's shoes, household cleaning products) was supported.

In order to determine whether systematic bias in perception of roles existed among couples who did not agree, Davis and Rigaux defined two types of nonconsensus, resulting from modesty or vanity (p. 58):

modesty--either or both spouses overestimating the other's influence in a decision or understanding their own influence.

vanity--either or both spouses overestimating their own influence in a decision or understanding the other's influence.

Davis and Rigaux utilized a 5 percentage point differential between aggregate modesty and vanity as an operational criterion for systematic bias in nonconsensus. They found a majority of cases indicating vanity. Among the 19 decision items common to both studies, 6 of 9 nonconsensus items in the problem recognition phase were vanity biased, all 10 in the search for information phase, and 7 of 9 in the final decision phase.

The vanity bias results were not supported in the present study. At the problem recognition phase, there was no particular bias for either the total sample (5 modesty and 6 vanity) or the traditional couples (5 each). While the total sample was in line with the earlier results at the search phase (8 of 9 vanity biased), traditional couples were more evenly divided (4 modesty and 5 vanity). Finally, at the final decision phase, the total sample was marked by more systematic bias cases (13 to 9) than the previous study as well as having much greater modesty bias (11 of the 13 cases) than was previously found. Since no consistent bias difference pattern emerged through inspection of the 3 decision items described the same in both studies (household cleaning products, children's toys for birthdays and holidays, and family vacation), it can be suggested these between-study differences in systematic nonconsensus bias were a result of differences in item specification.

There were also some differences between traditional couples and all couples with respect to nonconsensus bias. At the recognition phase, systematic bias in the modesty direction among traditional couples was found for washing machine and primary family car when there was no systematic bias among all couples for these items. Again, while there was no systematic bias among all couples for children's toys for birthdays and holidays and, to a lesser extent, household cleaning products, there was a vanity bias among traditional couples. Among all couples, a modesty bias was found for movies while a vanity bias was found among traditional couples.

At the second, search for information, phase, traditional couples were modesty biased with respect to primary family car and neck tie for the husband, while all couples were vanity biased in the former case and not systematically biased in the latter. Traditional couples, on the other hand, were vanity biased with respect to washing machine and children's toys for birthdays and holidays, while there was no systematic nonconsensus bias for the total sample for those decision items at this phase.

At the decision phase, the all couples group was found to be modesty biased with respect to sofa for living room or family room and replacement or additional pots and pans for the kitchen while the traditional couples exhibited no systematic nonconsensus bias. While there was no systematic nonconsensus bias for the total sample with respect to lawnmower, traditional couples exhibited a modesty bias.

TABLE 2
Extent of Husband-Wife Agreement About 20 Decisions
(In Percentage)

Key Number	Problem Recognition						Search for Information						Final Decision															
	Modesty		Consensus		Vanity		Modesty		Consensus		Vanity		Modesty		Consensus		Vanity											
	D/R	60/28	D/R	60/28	D/R	60/28	D/R	60/28	D/R	60/28	D/R	60/28	D/R	60/28	D/R	60/28	D/R	60/28										
1	24	12	15	57	68	63	19	20	22	17	13	7	69	67	68	14	20	25	14	22	21	80	68	61	6	10	18	
2	8	13	11	80	77	75	12	10	14	9	8	11	82	87	78	9	5	11	28	39	37	49	41	33	23	20	30	
3	9	10	7	86	83	82	5	7	11	7	12	14	83	78	72	10	10	14	14	23	25	72	64	54	14	13	21	
4	27	15	18	55	72	82	18	13	0	22	22	15	54	58	59	24	20	26	16	19	21	63	66	65	21	15	14	
5	10	3	4	74	89	92	16	8	4	9	3	7	71	85	82	20	12	11	17	15	19	61	73	77	22	12	4	
6	17	24	30	59	59	51	24	17	19	17	14	19	62	58	46	21	28	35	18	21	31	59	67	46	23	12	23	
7	24	17	29	53	63	64	23	20	7	18	17	21	58	61	68	24	22	11	19	25	19	64	58	62	17	17	19	
8	14	14	11	71	81	85	15	5	4	15	10	7	68	85	89	17	5	4	14	14	12	67	82	88	19	4	0	
9	22	20	14	64	60	75	14	20	11	25	20	18	52	60	71	23	20	11	24	23	21	54	55	58	22	22	21	
10	15	5	4	70	92	89	15	3	7	9	5	0	74	92	96	17	3	4	14	5	4	72	88	89	14	7	7	
11	7	4	0	86	85	92	7	11	8	10	2	0	83	91	100	7	7	0	17	5	8	63	82	84	20	13	8	
12	5	8	0	91	85	86	4	7	14	2	5	0	90	92	96	8	3	4	5	3	4	82	92	89	13	5	7	
13	15	16	8	70	66	63	15	18	29	22	20	16	59	60	56	19	20	28	18	14	8	60	65	64	22	21	28	
14	15	19	18	60	56	61	25	25	21	10	12	11	61	71	60	29	17	29	17	17	18	59	66	57	24	17	25	
15	8	10	11	79	68	75	13	22	14	5	8	7	81	77	82	14	15	11	18	23	29	68	60	60	14	17	11	
16	9	15	14	72	65	75	19	20	11	7	10	11	80	75	82	13	15	7	18	22	21	60	61	65	22	17	14	
17	13	20	25	64	70	64	23	10	11	4	18	21	70	64	65	26	18	14	9	20	25	68	65	57	23	15	18	
18	21	25	18	58	61	55	21	14	27	17	20	24	56	64	57	27	16	19	14	23	30	76	62	53	10	15	17	
19	18	24	19	64	61	62	18	15	19	14	21	19	61	61	66	25	18	15	5	15	19	88	75	77	7	10	4	
20	--	2	4	--	88	82	--	10	14	--	3	4	--	89	89	--	8	7	--	10	4	--	78	84	--	12	12	20
Average ^d	15	14	13	69	72	74	16	14	13	13	13	12	69	73	73	18	14	15	16	18	20	66	68	65	18	14	15	

a. Davis and Rigaux (1974) shown in percentages, 73 couples and 25 decisions.

b. 60 couples selected by means of a cluster sample.

c. 28 of the 60 couples classified as "traditional," that is, masculine husband and feminine wife.

d. 19 items common to both studies.

Product Classes and Systematic Nonconsensus
Perceptual Bias of Roles

It may be possible to generalize about perceptual bias in marital roles in terms of product classes. The classes used were convenience goods, shopping goods, specialty goods, and preference goods; a scheme suggested by Holbrook and Howard (1977) as cited by Frankel (undated). The standard definitions were utilized for the first three goods classifications, with preference goods defined as "those for which there is a strong brand preference, but low shopping effort is made (Frankel, p. 18)." It was recognized, of course, that the classification of any product is open to question since it would be a result of perceptions about how the consumer purchasing agents in question shop for those goods.

The convenience goods appeared to be beef roast and, somewhat surprisingly perhaps, adhesive bandages. These products were characterized by modesty bias at all phases. The preference goods appeared to be tooth paste, lawnmower, children's shoes, and replacement tires for the primary family car. Preference goods appear to be characterized by vanity bias, either in all 3 phases or a number of combinations of 2 phases, but not modesty biased in the recognition or final decision phase. The shopping goods appeared to be sofa for living room or family room, replacement of additional pots and pans for the kitchen, primary family car, insurance on husband's life, homeowner's or renter's insurance, movies, and family vacation. All these items were characterized by a modesty bias in the decision phase. Children's toys for birthdays and holidays appeared to be specialty goods with no systematic bias in the first 2 phases and a vanity bias in the final decision phase. Not all items were classifiable in this manner. Household cleaning products, neck tie for the husband, slacks for the wife, and washing machine were not characterized by systematic bias at any stage. These items could well have been inadequately specified as it was expected that individual household cleaning products would be preference goods while the remaining 3 would be shopping or specialty goods.

Among the traditional couples, children's toys for birthdays and holidays appeared to be a preference rather than a specialty item, while washing machine appeared to be a shopping good, and there was no clear pattern for replacement or additional pots and pans for the kitchen.

Conclusions and Implications

As in the Davis and Rigaux (1974) study, these results are based on rather intensive analysis of a small data base. These results are useful, however, in that a non-convenience sample was utilized permitting generalization to a specifiable population.

The general movement toward role specialization in search for information about alternatives found by Davis and Rigaux was not supported, as such, in this study. A major reason was due to the fact that problem recognition specialization was more prominent in this study than was true previously. Specialization at the problem recognition phase, in turn, was likely to be a result of the specific nature of the decision items rather than a difference in sample populations or methods.

Role specialization obviously does exist in household decision processes and the implications discussed by Davis and Rigaux (pp. 59, 60) have clearly not been mitigated by these results. These results strongly support, however, clearly specifying decision items in research. In fact, item attributes (e.g., color of

car) should be specified, although this information may be impossible to obtain for all decision process phases. Previous research (e.g., Davis, 1970; Hempel, 1972) has specified item attribute specialization, but not across decision process phases.

Davis and Rigaux held that one implication of high non-consensus in role perceptions was a need to word questions more specifically (p. 60) as was done in the present study. Interestingly enough, however, high item specificity appeared to have little effect on total consensus, 68% for the items common to both studies in the present study at the decision phase compared to 66% in the Davis and Rigaux study. Evidently communicators must take note of the degree and direction of nonconsensus rather than hoping to eliminate it through question wording. Nonetheless, the position taken throughout this analysis has been in favor of high specificity of items.

Decision makers should want to be knowledgeable about whether or not there is a vanity bias in perception of marital roles during the household decision process. It stands to reason that individual spouses with a vanity bias would get involved in the decision process regardless of role specialization, possibly in the phase not covered in these studies: evaluation of the final decision. These cases would thus require communication to both spouses.

Finally, while some differences were apparent between traditional couples and all couples in terms of role specialization at different phases of the decision process and level of role consensus, the number of these differences, while large enough to be significant, was also small enough to be, at least potentially, a result of random error. Thus, it would be difficult to make any generalizations with confidence. Nonetheless, the potential for differences appears established enough to be further studied in future research.

Appendix

Comparison of Economic Decision Items Utilized
In Present Study And Davis and Rigaux²

Key Number	Present Study	Davis and Rigaux
1	Sofa for living room or family room	Living room furniture
2	Drapes for living room	Other furnishings (rugs, drapes)
3	Replacement or additional pots and pans for the kitchen	Kitchenware
4	Washing machine	Household appliances excluding TV
5	Lawnmower	Garden tools
6	TV for living or family room	TV, Hi-Fi, tape recorder
7	Family car (primary)	Car
8	Beef roast	Food and nonalcoholic beverages
9	Neck tie for husband	Husband's clothes

Key Number	Present Study	Davis and Rigaux
10	Slacks for the wife	Wife's clothes
11	Children's shoes	Child(ren)'s clothes
12	Household cleaning products	Household cleaning products
13	Children's toys for birthdays and holidays	Child(ren)'s toys for birthdays and holidays
14	Tooth paste	Cosmetics and toiletries
15	Insurance on the husband's life	Life insurance
16	Homeowner's or renter's insurance	Other insurance
17	Adhesive bandages	Nonprescription drugs and first-aid items
18	Movies	Concerts, movies, theater
19	Family vacation	Family vacation
20	Replacement tires for the family car (primary)	No comparable item

^aDavis and Rigaux (1974)

References

S. L. Bem, "The Measurement of Psychological Androgyny," Journal of Consulting and Clinical Psychology, 42 (April 1974), 155-62.

S. L. Bem, "Sex Role Adaptability: One Consequence of Psychological Androgyny," Journal of Personality and Social Psychology, 31 (April 1975), 634-43.

H. L. Davis, "Dimensions of Marital Roles in Consumer Decision-Making," Journal of Marketing Research, 7 (May 1970), 168-77.

H. L. Davis, "Measurement of Husband-Wife Influence in Consumer Purchase Decisions," Journal of Marketing Research, 8 (August 1971), 305-12.

H. L. Davis and B. P. Rigaux, "Perception of Marital Roles in Decision Processes," Journal of Consumer Research, 1 (June 1974), 51-62.

M. Frankel, A Summary Report: What Do We Know About Consumer Behavior?, Washington, D. C.: National Science Foundation, undated.

D. H. Granbois, and R. P. Willett, "Equivalence of Family Role Measures Based on Husband and Wife Data," Journal of Marriage and the Family, 32 (February 1970), 68-72.

D. J. Hempel, "A Cross-Cultural Analysis of Husband-Wife Roles in House Purchase Decisions," in M. Venkatesan, Proceedings, Association for Consumer Research, Chicago: ACR, 1972, 816-29.

M. L. Holbrook and J. A. Howard, "Frequently Purchased Goods and Services," in R. Ferber, A Synthesis of Selected Aspects of Consumer Behavior, Washington, D. C.:

National Science Foundation, 1977.

S. Sudman, Applied Sampling, (New York: Academic Press, 1976).

J. Scanzoni, "A Note on the Sufficiency of Wife Responses in Family Research," Pacific Sociological Review, 8 (Fall 1965), 109-15.

E. A. Wilkening and D. E. Morrison, "A Comparison of Husband and Wife Responses Concerning Who Makes Farm and Home Decisions," Marriage and Family Living, 25 (August 1963), 349-51.

D. M. Wolfe, "Power and Authority in the Family," in D. Cartwright (ed.), Studies in Social Power, (Ann Arbor: University of Michigan Press, 1959).

AN EXAMINATION OF
INNOVATIVE COMMUNICATORS, OPINION LEADERS AND INNOVATORS
FOR MEN'S FASHION APPAREL

Elizabeth C. Hirschman, University of Pittsburgh
William O. Adcock, Georgia State University

Abstract

Recent theoretical and empirical papers have suggested a need for more meaningful and useful adopter categories. The findings in this study provide support for the construct validity and pragmatic utility of an adopter category composed of individuals who are above average in both innovativeness and opinion leadership. A comparison of survey results from two socioeconomically diverse markets points to the potential danger of generalizing from research results obtained within a limited social environment.

Introduction

Diffusion of innovations is one of the most widely researched topics in the behavioral sciences. Rogers and Shoemaker (1971) compiled a bibliography of over 1,000 diffusion research studies which had been conducted in thirteen separate disciplines. Major diffusion research traditions exist in seven of these disciplines: anthropology, early sociology, rural sociology, education, medical sociology, communication, and marketing.

Despite the massive, combined research effort represented by the existing body of diffusion research, many observers are critical of the methodology, assumptions, and concepts associated with such studies, and many doubt that results are meaningful or useful. Rogers (1976) states that most diffusion research is lacking in originality and that contributions to more effective social programs have been few.

Taylor (1977) discusses the frustration felt by many consumer researchers that no meaningful set of relationships has yet been identified to characterize early triers of new products. Kotler and Zaltman (1976) note that research into early adopter characteristics has produced many contradictory results, and that few studies have proven useful in marketing planning. They also state that little information has been generated with regard to the sociological and psychological variables influencing the adoption decision.

Recently, Rogers (1976) has commented on the need for improved research designs, including network analysis and longitudinal research studies. He also stresses that scholars have accepted the diffusion paradigm so completely that they may be unable to recognize or to overcome possible conceptual biases.

The authors of this study concur with these scholars that diffusion research results have often been confusing and of limited usefulness, particularly in the marketing discipline. Improved research designs may lead to more meaningful results in the future. However, it appears that some of the basic concepts that have been widely accepted in diffusion research also need to be reexamined as a prelude to further analysis. This paper is intended to enlarge upon several recent theoretical and empirical efforts in determining relevant adopter categories.

Categories in Diffusion Theory

Two categorizations are generally accepted as valid in diffusion theory. First, adoption patterns are assumed to follow a normal, bell-shaped distribution. Within this distribution five basic adopter categories are posited to exist: innovators (first 2.5 percent of population to adopt), early adopters (12.5 percent), early majority (34 percent), late majority (34 percent) and laggards (final 16 percent). The proportion of the population to be allocated to each adopter category reflects standard deviations within the normal distribution.

A number of marketing scholars have questioned the preciseness of this set of assumptions within a marketing context. Peterson (1973) states that a distribution in a marketing situation would probably be highly skewed, not normally distributed, because of the effects of increased numbers of producers and increased promotion as a new product began to win consumer acceptance. Peterson also proposes the use of algorithms to identify categories in specific product classes in such a way as to minimize variance within groups. The use of such an algorithm could result in more or less than five categories in any given situation. Other marketing researchers have questioned the percentage distributions within categories. Robertson (1971) favors a ten percent allocation to the innovator category in marketing studies, while King (1963) classified 35 percent of his sample as "early buyers."

The second major categorization paradigm used in adoption theory concerns personal influence. A basic assumption of diffusion theory is that some individuals within a given social structure are influential in persuading others to adopt products. Attempts to identify the characteristics of these opinion leaders, however, have produced contradictory and inconclusive results (Zaltman and Stiff, 1973).

Overlap Between Innovativeness and Opinion Leadership

The overlap between these two sets of diffusion categories, innovativeness and opinion leadership, would intuitively appear to hold a great deal of potential for increasing understanding of the diffusion process. Yet surprisingly little empirical research has been specifically addressed to this issue. Robertson (1971) tabulated results from twenty-one marketing studies which had been conducted between 1954 and 1969, and reported that thirteen attempted to study relationships between innovativeness and opinion leadership. He determined that ten found a positive relationship, while three found no relationship. In no study was a negative relationship found.

Several recent studies of women's fashion adoption have reported disappointing results relative to correlation between innovativeness and opinion leadership. King (1963) reported that early and late buyers were not significantly different with respect to influence. Studies by Robertson and Myers (1969) and by Myers and Robertson (1972) found overlap between innovativeness and opinion leadership; but in each instance the over-

lap was described as moderate or small.

In a study of male fashion adoption Darden and Reynolds (1974) identified six adopter groups, as opposed to the five groups in classic diffusion theory. Discriminant analysis showed that opinion leadership was strongly associated with the most innovative adopter groups.

The authors were able to locate only two studies within the marketing discipline which were designed specifically to look at individuals who were above average in both innovativeness and opinion leadership. The first study was conducted by Summers (1971). He examined the overlap between innovativeness and opinion leadership found in generalized change agents for women's fashion apparel. He concluded that the relationship was not strong, although it was statistically significant. Baumgarten (1975) coined the term for these generalized change agents in a study he conducted on adoption processes for men's fashion apparel. Calling this adopter category, innovative communicators, Baumgarten concluded that they were differentiable from the general population and developed a psychographic profile of their characteristics.

The following year Midgley (1976) made an important theoretical contribution to the diffusion theory paradigm by formulating a three dimensional set of adopter categories. The three dimensions he proposes are: trial, acceptance/rejection and opinion leadership. His model incorporates four categories. Active adopters are individuals who have tried an innovation and will provide favorable information on it. Active rejectors are individuals who have tried an innovation, found it unsatisfactory, and will provide unfavorable information on it. Passives are those who have tried the innovation, but who will not provide information on it. The fourth group is termed potential adopters and includes those who have not yet tried the innovation.

Kotler and Zaltman (1976) added further sophistication to the issue of innovator-opinion leader overlap, when they formulated a theory of the best prospect. Noting that early adopters may not necessarily be the best target market segment for a new product, they provide a cost-benefit equation for determining the value of a prospect. The four factors that enter into this equation include (1) propensity to innovate, (2) amount the prospect is likely to purchase, (3) additional purchases that the prospect is likely to stimulate through interpersonal influence, and (4) the cost of communicating with the prospect.

The Kotler-Zaltman model does not require that the "best prospect" be high in both innovativeness and opinion leadership. Instead, the key is the nature of the interaction between the two characteristics, with purchase probabilities and communication costs also important to the economics of a target market strategy.

In summary, some empirical research has been conducted on the overlap between innovativeness and opinion leadership. Generally, these studies have identified relationships possessing statistical significance, but lacking in pragmatic utility. The theoretical work by Midgley and Kotler and Zaltman and the empirical work of Summers and Baumgarten, however, appear to provide a basis for defining a new, and pragmatically valuable, early adopter category - the innovative communicator.

Purpose

The study reported here attempts to enlarge upon the earlier findings of Baumgarten (1975), with regard to male fashion "innovative communicators." These findings are certainly indicative of the possible existence and

potential utility of the innovative communicator construct. However, some methodological limitations of Baumgarten's study prohibit the development of a generalized theory concerning the role of innovative communicators in the fashion apparel adoption process. Specifically, his data were limited to a sample of 389 unmarried male undergraduates at Purdue University. Thus, there is the potential for a restriction of range for the findings. Further, because data were gathered in only one setting, there can be no assurance as to their external validity. The necessity of assessing the external validity of findings from a study of this type is of central importance to the development of consumer behavior theory. As Shocker and Zaltman (1976) state:

In the marketing literature on consumer research, one still finds relatively little explicit concern with the internal and external validity attending the research itself...Issues of validation are very central to the development of theory and to the progression of research from mere ad hoc responses to specific inquiries toward the cohesive bodies of knowledge characteristic of disciplines.

A further limitation of Baumgarten's study was its failure to examine other possible adopter categories; all respondents who were not categorized as innovative communicators were combined into an "others" category. The study reported here uses adoption/influence patterns to specify four adopter categories. Innovative communicators, as defined in this study and in the Baumgarten study, are individuals who rank high on both innovativeness and opinion leadership. A second group in this study, opinion leaders, includes respondents who score high on opinion leadership, but not on innovativeness. A third group, innovators includes individuals who score high on innovativeness, but not on opinion leadership. Finally, the general population includes those who do not score high on either construct.

The study reported here will also attempt to provide some preliminary examination of the four elements of the Kotler-Zaltman theory of the best prospect choice. In addition to measuring the interaction effect of overlapping innovativeness and opinion leadership, some data were gathered concerning media selection, store attribute evaluations and purchasing patterns.

Data Collection

Sampling Sites

Data for the research were gathered in two Southeastern cities during Spring, 1977. The cities and their respective sample sizes are given below:

Atlanta, Georgia	400
Augusta, Georgia	239
Total	639

The cities and sample sizes were specified by the retail firm sponsoring the research. In Table 1 are given a set of statistics comparing the two cities on a variety of attributes.

TABLE 1

	A Statistical Comparison of the Two Sample Cities	
	Atlanta	Augusta
Population	1,841,200	278,700
Average household income	\$15,162	\$12,201
Annual retail sales (000's)	\$4,791,964	\$752,957
1975-1980 pop. growth rates	10.1%	5.3%

SOURCE: 1976 Survey of Buying Power, Sales & Marketing Management Magazine.

Although the two cities are geographically adjacent, approximately 120 miles apart, they exhibit some large economic and demographic dissimilarities. The cities may be dichotomized using the statistics given in Table 1. At one end would be placed Atlanta, which is the larger, wealthier and more rapidly growing of the two cities. At the other end is Augusta with the smaller population, lower income per household and the slower rate of growth. This diversity between the cities comprising the sample should help provide a rich ground for assessing the external validity of the findings from this research.

Sampling Procedure

To obtain the sample, a list of all telephone exchanges within each city's S.M.S.A. was compiled. By working closely with telephone company executives in each city, a recent (within the past two weeks) estimate was obtained of the number of telephone-equipped households within each exchange. The overall sample collected from each city was then apportioned according to the number of telephone households falling within a particular exchange.

A random digit dialing technique was used to gather the sample, which eliminated possible bias against new arrivals and unlisted numbers which can be caused by using the telephone directory. All interviewing was conducted from 6:00 PM to 9:30 PM weeknights to help insure an adequate representation of working persons. Trained interviewers employed by commercial field survey research firms were used to conduct the interviews. In Table 2 are given the dates of interviewing for each city, the response rates, and the percentage of questionnaires which were validated.

TABLE 2

Sampling Statistics		
	Atlanta	Augusta
Response Rate*	80.6	83.4
Percent Validated	15.0	20.0
Interviewing Dates	2/7-2/27 (1977)	2/1-2/10 (1977)

*Once interview was initiated.

Criterion Operationalization

One of the problems inherent in past diffusion research is that innovativeness and opinion leadership have been operationalized inconsistently. Rogers (1976), Kotler and Zaltman (1976), and Zaltman and Stiff (1973), among others, have criticized the measurement techniques used in most studies. These authors have advocated the use of network analysis to more accurately measure the constructs under study. The authors of this paper agree that such research designs are most desirable. As a practical matter, however, the resources needed for such studies may be unavailable for many researchers. As in the case of the study reported here, more limited measurement and operational schemes will continue to be used, but researchers should attempt to select designs that offer reasonable hope of construct, as well as face, validity.

The study reported in this paper utilized sets of self-designating questions to measure both fashion innovativeness and opinion leadership. The two sets of questions, three for each construct, are listed in Table 3. Responses to the question set for each construct were summed to create a composite score.

TABLE 3

Measurement of Innovativeness and Opinion Leadership	
Fashion Innovativeness = Question A + Question B - Question C. Fashion Opinion Leadership = Question D + Question E + Question F.	
A. Are you willing to try new ideas about clothing fashions? How often?	4. Often 3. Sometimes 2. Seldom 1. Never 0. Don't Know
B. Do you try something new in the next season's fashions? How often?	4. Often 3. Sometimes 2. Seldom 1. Never 0. Don't Know
C. Are you usually among the last to try new clothing fashions? How often?	4. Often 3. Sometimes 2. Seldom 1. Never 0. Don't Know
D. How often do you influence the types of clothing fashions your friends buy?	4. Often 3. Sometimes 2. Seldom 1. Never 0. Don't Know
E. How often do others turn to you for advice on fashion and clothing?	4. Often 3. Sometimes 2. Seldom 1. Never 0. Don't Know
F. How many of your friends and neighbors regard you as a good source of advice on clothing fashions?	4. Almost everyone I know 3. More than half 2. Less than half 1. Almost no one 0. Don't Know

Male Fashion Innovative Communicators were operationally defined in this research as men who scored one standard deviation above the mean on both fashion innovativeness and fashion opinion leadership. These persons composed 6.8% of the sample (n = 16) in Augusta and 7.8% of the sample (n = 31) in Atlanta.

To describe the differentiating characteristics of the Male Fashion Innovative Communicator, consideration must be given to the groups against which it is appropriate to make comparisons. It was felt that for the purposes of assessing construct validity three relevant groups would be required. These three groups consisted of (1) Fashion Innovators - operationalized as persons scoring one standard deviation above the mean on fashion opinion leadership. (2) Fashion Opinion Leaders - operationalized as persons scoring one standard deviation above the mean on fashion opinion leadership but below one standard deviation above the mean on fashion innovativeness and (3) the General Population operationalized as persons scoring below one standard deviation above the mean on both fashion innovativeness and fashion opinion leadership.

The distributions obtained for the four groups are given in Table 4.

TABLE 4

	Construct Distributions			
	Atlanta		Augusta	
	n	%	n	%
Innovative Communicators	31	7.8	16	6.8
Innovators	30	7.5	19	8.1
Opinion Leaders	42	10.5	27	11.4
General Population	297	74.3	174	73.7

Findings

The variables used to compare male fashion innovative communicators with the three other groups of interest fell into five categories: (1) socioeconomic characteristics; (2) sociographic influences; (3) evaluations of retail store attributes; (4) media usage; and (5) store patronage.

The Atlanta sample received primary attention in the analysis reported here, rather than Augusta, because fashion interest was expected to be greater in the more cosmopolitan of the two cities. The analysis for each

category of variables involved three major steps:

1. First, the data were examined to identify differences between the general population and the three construct groups (male fashion innovative communicators, male fashion opinion leaders, and male fashion innovators).
2. Next, the data were further examined to determine if differences existed among the three construct groups. Identification of significant differences would provide some indication of construct validity for the innovative communicator adopter category.
3. Finally, results from the Atlanta sample were compared with those in Augusta to test their external validity.

In each instance, the Chi Square statistic was used to test the significance of differences among groups (cut off at .10 level).

Socioeconomic Characteristics

The variables examined included income, age, education, and length of residence in present city. Data reported in Table 5 shows that significant differences exist between the general population and the three construct groups only with respect to age. All three construct groups tend to include higher proportions of younger persons. This finding was also reflected in data from Augusta (Table 9). However, in neither Atlanta or Augusta were any significant differences found among the three construct groups with respect to age or any other socioeconomic variables.

Sociographic Influences

Variables relating to participation in social activities and to memberships in groups were expected to provide valuable insights into the life styles and information transfer characteristics of male fashion adopters.

Social Activities. As shown in Table 6, all three construct groups are significantly different from the general population with regard to participation in a diverse group of activities. The construct groups are more likely than the general population to include individuals who entertain at home, attend movies, and attend concerts. Similar results were reflected by data from the Augusta sample.

The data in Table 6 also suggest a systematic and consistent pattern of differences among innovative communicators and other two construct groups. Although statistically significant differences exist for only four of the eight activities (attending concerts, gardening, golfing, and camping), the innovative communicator group contains the smallest proportion of non-participants for each of the activities reported here. This is true even though other groups contain higher proportions of very active participants for some activities. Thus, it appears that innovative communicators participate in a wide range of social activities, although levels of commitment to some activities may not be high. The Augusta data, however, do not reflect similar patterns of differences among the construct groups.

This finding may be attributed to three possible factors: (1) the small size of the Augusta sample introduces sampling error; (2) the pattern of correlations found in the Atlanta data are spurious, or (3) the three constructs manifest themselves differently in various social settings. Atlanta is a large, wealthy and highly transient city relative to Augusta; it is possible that the great variance in social structure

between these two cities may act to influence the form and function of social phenomena such as innovative communication, innovativeness and opinion leadership. Further, the types of and functions served by social activities and group memberships in the two cities may also be radically different.

If this third factor is indeed the most accurate rationale for the differences observed in the data, a strong argument could be made for greater attention being given to the impact of sociological variables on the adoption process, rather than limiting focus only to psychological correlates.

Group Membership. Respondents were asked to name groups and organizations to which they belong. The results indicate that unexpected behavior exists on the part of opinion leaders. Members of this group are significantly less likely than members of the other two construct groups to belong to social and recreational groups or organizations. Apparently, these individuals exercise personal influence within a less formalized or structured environment than might have been anticipated. Similar membership patterns are suggested by the Augusta data, but the results do not achieve statistical significance. Again, the impact of the social environment upon the manifestation of the construct is strongly suggested by the data.

Store Attribute Evaluations

Respondents were asked if each of ten store attributes was of much concern, moderate concern, or little concern to them in deciding where to shop. A comparison of evaluations of store attributes among the three construct groups and the general population is felt to have potential value for the development of more effective retail marketing strategies. One of the central criticisms of past diffusion research is that it has failed to demonstrate pragmatic utility (Kotler and Zaltman, 1976).

One finding in this area may be of value to both retail strategists and to consumer behavior theorists. As shown in Table 7, significant differences (.10 level) exist among the three construct groups for only one attribute, the store's guarantee, exchange and adjustment policies. The groups that exhibit innovative tendencies (innovative communicators and innovators) express more concern for guarantee, exchange and adjustment policies than do opinion leaders. Similar results, also significant at the .10 level, were found in Augusta.

This finding may indicate that an emphasis on guarantees or adjustment policies represents a form of risk reduction behavior on the part of male fashion innovators and innovative communicators. This would appear quite logical from the point of view of the fashion innovative consumer. Purchasing an unusually styled garment or one constructed from out-of-the-ordinary materials (for example, feather, fur or sequins) may activate both the social and economic components of perceived risk. Thus, if the store has a well-thought-of policy for merchandise return, exchange or repair, this may be an especially valuable attribute to the innovator.

Media Usage

Data were gathered only about newspaper subscriptions and radio station listenership, since these two media are of most pragmatic relevance from a retailing standpoint.

Data about newspaper readership included overall sub-

scription rates and preferences for morning versus afternoon newspapers. In Atlanta, neither set of data identified differential tendencies toward newspaper readership among the four groups of interest in this study, (Table 8). However, in Augusta there were significant differences (.05) in subscription rates to the morning versus afternoon newspapers. These differences existed for both the three construct groups and the general population and among the three construct groups themselves. In Augusta the innovative communicators and opinion leaders were much more likely to subscribe to the afternoon newspaper than were innovators or the general population.

There are also significant and intuitively appealing differences among the four groups in terms of radio station preference. In both Atlanta and Augusta, the three construct groups display a distinct preference for stations featuring a progressive rock or jazz format, while the general population tends to be more partial to stations with classical or easy-listening type music. Differences among the four groups were significant at the .01 level in both Atlanta and Augusta. However, the Atlanta data do not reflect any significant differences among the three construct groups; while the Augusta data do.

Store Patronage

It was expected that the four groups in question would display differential patronage patterns for the purchase of apparel. Respondents were asked where their last purchase had been made for mens suits and mens sportswear. Their responses were categorized into specialty and traditional department stores, national chain department stores, and discount stores.

Contrary to expectations, no significant differences among groups were found. It is possible that the "last purchase" specification was inadequate as an operationalization for patronage and that "usual store" would have provided somewhat different results. It is also possible that the store categories were inappropriate. In any event, survey data provide no basis for concluding that adopter groups prefer different types of stores.

Summary of Findings

The data discussed here identify significant patterns of differences at three levels of analysis: between the general population and construct groups; among construct groups; and between two socio-economically diverse markets.

Data show that Innovative Communicators, Opinion Leaders and Innovators tend to be younger than the general population, more likely to participate in several social activities, and prefer jazz/rock-oriented radio stations.

More importantly, significant differences were identified among the three construct groups. Innovative communicators are least likely to be non-participants in a wide range of activities. Opinion leaders are least likely to be members of social or recreational groups. Innovative communicators and innovators are more likely to express concern about guarantees and store adjustment policies than are opinion leaders. Findings of significant differences indicate that each group is unique and that the constructs possess more than face validity. It is also important to recognize that some findings in Atlanta were not replicated in Augusta. Differences in results may reflect the differences in the social structures and economic environments of the two cities. These differences suggest that it may be an error to attempt to generalize too widely from a single study

pertaining to a limited social environment. Unless results are tested for external validity, generalizations are probably not justified.

This study provides empirical support for the Kotler-Zaltman theory of the best prospect choice. Since male fashion innovative communicators are above average in both opinion leadership and innovativeness and also express differential patterns of media preference, it should be possible to develop communications strategies directed toward these potential best prospects. However, further research will be needed to determine if shopping patterns and purchase propensities differentiate this group from other consumers.

Implications

The study reported here is intended to be transitional. The diffusion paradigm clearly needs to be examined from fresh perspectives and reformulated in a way that will be more useful to scholars in various disciplines as well as to marketing practitioners. Any new formulation of diffusion theory should involve a redefinition of adopter categories.

Findings that two unique groups of individuals, innovative communicators and innovators, exhibit innovative behavior help to explain the failure of earlier research to identify meaningful correlates of such behavior. A single set of predictor variables would be unlikely to describe members of both groups. For the same reason, findings that two unique groups of individuals, innovative communicators and opinion leaders, exercise personal influence explain the inconsistent results of studies intended to identify correlates for opinion leadership.

Findings from this study also suggest that failure to take sociological factors into account and failure to test for external validity have caused much of the previous research into innovation and opinion leadership to be of limited usefulness.

This study is not intended to be a definitive statement as to the nature of the adopter categories which will eventually prove to be most useful. The value of the findings reported here may simply be to suggest reasons why other studies have produced few useful insights and to demonstrate that new categories can be specified. Additional research involving network analysis and longitudinal designs may be needed to go beyond the two-dimensional scheme illustrated here, but such efforts are desirable if diffusion theory is to be made more meaningful.

References

- Steven A. Baumgarten, "The Innovative Communicator in the Diffusion Process," Journal of Marketing Research, 12 (February 1975), 12-18.
- William R. Darden and Fred D. Reynolds, "Backward Profiling of Male Innovators," Journal of Marketing Research, 11 (February 1974), 79-85.
- Charles W. King, "Fashion Adoption: A Rebuttal to the 'Trickle Down' Theory," in Stephen A. Greyser, editor, Toward Scientific Marketing, Chicago: American Marketing Association, 1963, 108-125.
- Philip Kotler and Gerald Zaltman, "Targeting Prospects for a New Product," Journal of Advertising Research, 16 (February 1976), 7-18.

David F. Midgley, "A Simple Mathematical Theory of Innovative Behavior," Journal of Consumer Research, 3 (June 1976), 31-41.

James H. Myers and Thomas S. Robertson, "Dimensions of Opinion Leadership," Journal of Marketing Research, 9 (February 1972) 41-46.

Robert A. Peterson, "A Note on Optimal Adopter Category Determination," Journal of Marketing Research, 10 (August 1973) 325-9.

Thomas S. Robertson, Innovative Behavior and Communication, New York: Holt, Rinehart and Winston, Inc., 1971.

Thomas S. Robertson and James H. Myers, "Personality Correlates of Opinion Leadership and Innovative Buying Behavior," Journal of Marketing Research, 6 (May 1969), 164-8.

Everett M. Rogers and Floyd Shoemaker, Communication of Innovations, New York: Free Press, 1971.

Everett M. Rogers, "New Product Adoption and Diffusion," Journal of Consumer Research, 2 (March 1976), 290-301.

Allen D. Shocker and Gerald Zaltman, "Validity Importance in Consumer Research: Some Pragmatic Issues," in William D. Perreault, Jr., editor, Advances in Consumer Research, Volume IV, Atlanta: Association for Consumer Research, 1977, 405-8.

John O. Summers, "Generalized Change Agents and Innovativeness," Journal of Marketing Research, 8 (August 1971), 313-6.

James W. Taylor, "A Striking Characteristic of Innovators," Journal of Marketing Research, 14 (February 1977), 104-107.

Gerald Zaltman and Ronald Stiff, "Theories of Diffusion," in Scott Ward and Thomas S. Robertson, editors, Consumer Behavior: Theoretical Sources, Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1973, 416-468.

TABLE 5
Socioeconomic Characteristics
(Based Upon Atlanta Sample)

	(1) INNOV. COMMUN. (%)	(2) OPINION LEADER (%)	(3) INNOVATOR (%)	(4) GENERAL POPULATION (%)	SIGNIFICANCE	
					AMONG 4 GROUPS	AMONG 3 GROUPS
INCOME:						
UNDER \$7,500	9.7	14.3	6.7	8.1		
\$7,500 - \$14,999	54.7	42.9	35.5	48.5		
\$15,000 OR MORE	35.6	42.8	56.7	43.4		
TOTAL	100.0	100.0	100.0	100.0	N.S.	N.S.
RESPONDENT'S AGE						
18 - 34	74.2	66.7	60.0	42.7		
35 - 54	22.6	26.2	30.0	37.4		
55 OR OLDER	2.2	7.1	10.0	19.6		
TOTAL	100.0	100.0	100.0	100.0	.02	N.S.
RESPONDENT'S EDUCATION						
NOT H.S. GRADUATE	16.1	11.9	10.0	16.9		
H.S. GRADUATE	19.4	33.3	30.0	33.0		
SOME COLLEGE	29.0	33.3	20.0	19.2		
COLLEGE GRADUATE	35.5	21.5	40.0	30.9		
TOTAL	100.0	100.0	100.0	100.0	N.S.	N.S.
LENGTH OF TIME IN ATLANTA						
3 YEARS OR LESS	32.3	16.7	16.7	16.5		
3 - 6 YEARS	16.1	14.3	20.0	14.8		
OVER 6 YEARS	51.6	69.0	53.3	68.7		
TOTAL	100.0	100.0	100.0	100.0	N.S.	N.S.
SAMPLE SIZE	(31)	(42)	(30)	(297)		

TABLE 6

Sociographic Variables
(Based Upon Atlanta Sample)

	(1) INNOV. COMMUN. (%)	(2) OPINION LEADER (%)	(3) INNOVATOR (%)	(4) GENERAL POPULATION (%)	SIGNIFICANCE	
					AMONG 4 GROUPS	AMONG 3 GROUPS
SOCIAL ACTIVITIES						
ENTERTAIN AT HOME - VERY OFTEN	48.4	28.6	43.3	16.2		
NEVER	6.5	14.3	16.7	20.3	.01	N.S.
SPECTATOR SPORTS - VERY OFTEN	35.5	35.7	33.3	22.3		
NEVER	6.5	9.5	23.3	20.3	N.S.	N.S.
MOVIES - VERY OFTEN	29.0	19.0	33.3	9.5		
NEVER	3.2	9.5	13.3	26.0	.01	N.S.
CONCERTS - VERY OFTEN	25.8	14.3	13.3	4.4		
NEVER	19.4	45.2	56.7	61.5	.01	.05
GARDEN - VERY OFTEN	25.8	21.4	26.7	27.4		
NEVER	29.0	54.8	46.7	34.8	.02	.10
GOLF - VERY OFTEN	22.6	7.1	6.7	5.7		
NEVER	45.2	71.4	73.3	78.0	.01	.05
TENNIS - VERY OFTEN	16.1	16.7	20.0	12.2		
NEVER	35.5	52.4	53.3	60.8	N.S.	N.S.
CAMP - VERY OFTEN	16.1	11.9	13.3	11.8		
NEVER	38.7	47.6	66.7	58.4	.08	.10
MEMBERSHIP IN GROUPS OR ORGANIZATIONS						
SOCIAL	25.8	11.9	33.3	11.8	.01	.10
RECREATIONAL	25.8	4.8	16.7	8.8	.01	.05
RELIGIOUS	22.6	23.8	20.0	34.0	.10	N.S.
BUSINESS	19.4	14.3	16.7	19.2	N.S.	N.S.
COMMUNITY	12.9	11.9	26.7	15.5	.10	N.S.

TABLE 7
Concern For Store Attributes
(Based Upon Atlanta Sample)

	(1)	(2)	(3)	(4)	SIGNIFICANCE	
	INNOV. COMMUN. (%)	OPINION LEADER (%)	INNOVATOR (%)	GENERAL POPULATION (%)	AMONG 4 GROUPS	AMONG 3 GROUPS
PERCENTAGE EXPRESSING "MUCH CONCERN"						
GUARANTEES/ADJUSTMENT POLICY	96.9	75.6	86.7	83.0	N.S.	.10
QUALITY OF MERCHANDISE	89.7	82.5	90.0	86.6	N.S.	N.S.
PRICES	78.6	71.4	56.7	74.7	N.S.	N.S.
SALES REPRESENT SAVINGS	78.4	60.0	73.3	67.7	N.S.	N.S.
SALES CLERK SERVICE	69.0	63.4	55.2	54.7	N.S.	N.S.
LOCATION	69.0	56.1	70.0	67.1	N.S.	N.S.
VARIETY	65.5	59.5	73.3	68.7	N.S.	N.S.
LAYOUT	42.9	39.0	40.0	36.9	N.S.	N.S.
CREDIT AND BILLING POLICIES	37.9	25.6	34.5	28.5	N.S.	N.S.
DISPLAYS	37.9	32.5	36.7	38.8	N.S.	N.S.

TABLE 8
Media Usage Patterns
(Based Upon Atlanta Sample)

	(1)	(2)	(3)	(4)	SIGNIFICANCE	
	INNOV. COMMUN. (%)	OPINION LEADER (%)	INNOVATOR (%)	GENERAL POPULATION (%)	AMONG 4 GROUPS	AMONG 3 GROUPS
NEWSPAPER SUBSCRIPTIONS:						
MORNING	36.8	44.4	26.7	36.9		
EVENING	31.6	37.0	53.3	45.5	N.S.	N.S.
RADIO PREFERENCE:						
JAZZ/ROCK	62.9	54.2	56.3	21.6		
CLASSICAL/EASY LISTENING	37.1	45.8	43.7	78.4	.01	N.S.

TABLE 9
Correlates Found Significant In Atlanta,
Tested For External Validity In Augusta

	(1)	(2)	(3)	(4)	SIGNIFICANCE	
	INNOV. COMMUN. (%)	OPINION LEADER (%)	INNOVATOR (%)	GENERAL POPULATION (%)	AMONG 4 GROUPS	AMONG 3 GROUPS
<u>Demographic Characteristics</u>						
Respondent's Age					.01	N.S.
10 - 34	87.6	77.8	73.7	44.8		
35 - 54	6.2	11.1	26.3	29.3		
55 or Older	6.2	11.1		25.9		
<u>Social Activities</u>						
Entertain at Home - Very Often	53.3	33.3	52.6	18.5		
Never	6.7	18.5	10.5	28.9	.01	N.S.
Movies - Very Often	46.7	44.4	36.8	13.2		
Never	13.3	3.7	10.5	31.0	.01	N.S.
Concerts - Very Often	13.3	22.2	15.8	4.7		
Never	53.3	55.6	73.7	66.7	.10	N.S.
Garden - Very Often	37.5	18.5	26.3	32.2		
Never	31.3	37.0	36.8	29.3	N.S.	N.S.
Camp - Very Often	6.7	48.1	15.8	20.1	.01	.01
Never	40.0	33.3	26.3	51.1		
Golf - Very Often	6.7	11.1	10.5	14.5		
Never	73.3	55.6	78.9	71.5	N.S.	N.S.
<u>Memberships</u>						
Social	12.5		10.5	10.9	N.S.	N.S.
Recreational	12.5	11.1	20.1	9.2	N.S.	N.S.
Community	12.5	14.8	5.3	12.1	N.S.	N.S.
<u>Store Attribute</u>						
Guarantees/Adjustment Policy	87.5	76.9	94.7	81.4	.10	N.S.
<u>Newspaper Subscriptions</u>						
Morning	85.7	90.0	41.7	75.8		
Evening	14.3	10.0	58.3	25.2	.05	.05
<u>Radio Preference</u>						
Jazz/Rock	92.8	52.9	62.5	41.5		
Classical/Easy Listening	7.2	47.1	37.5	58.5	.01	.01
Sample Size	(16)	(27)	(19)	(174)		

INDIVIDUAL TIME ORIENTATION AND CONSUMER LIFE STYLE

Robert B. Settle
Pamela L. Alreck
John W. Glasheen
San Diego State University

Abstract

F-A-S-T, a four-dimensional, trait-specific personality test was used to measure the time orientations of 599 adult subjects. These were compared to their responses to 406 Activity, Interest and Opinion items. The study reveals clearly discernible translations of individual time orientation into differing life style patterns.

Introduction

Of the wide variety of concepts and variables used by researchers to explain consumer behavior, personality factors have provided some of the most disappointing results. Yet despite the negative findings, we continue to pursue these illusive relationships because the concepts of personality have great intuitive appeal. Common observations of everyday life reveal persistent individual differences, earmarking specific personality types or traits, in both ourselves and in others. Judging from the literature on personality and buyer behavior, however, the role of personality in determining behavior in this area must be miniscule, indeed. Kassarian (1971) cites four possible causes for the failure of personality studies to explain variation in buyer behavior: (1) Few studies have been designed to test specific hypotheses obtained from a theoretical base, (2) we often expect the influence of personality variables to explain too large a portion of the variance in consumer behavior, (3) the instruments used are seldom designed to measure traits that could be expected to directly influence consumer decisions, and (4) the tests used are often standardized on special groups, rather than consumers in general. We agree.

This study has been designed to assess the role of one aspect of personality in influencing consumer behavior, while avoiding the pitfalls cited above. A point by point consideration, without extraneous detail, may clarify the intent of the project. First, the theoretical justification for the study arose from the belief that a holistic approach to the understanding of consumer behavior is ultimately more productive than a reductive-functional view (Markin, 1974, pp. 62-76). We were, then, more interested in discovering the relationship between a set of personality traits and consumer behavior patterns, rather than the selection of Brand X as opposed to Brand Y.

The typical general model of consumer behavior comprehends buyer influences that can be divided into three basic groups: Those relating to the various stimulus factors, those which affect the situation, setting or circumstances in which the behavior is performed, and those that fall under the rubric of personal factors. Among the latter, "personality" is but one of many very potent variables. Beyond this fact, personality, itself, can be divided into a myriad of traits and dimensions. Consequently, we sought traits that might be major influences, but were determined to be satisfied with broad patterns of relationships and relatively small proportions of explained variance.

The instruments used in the study were designed to measure traits that can very definitely be expected to

directly affect consumer decisions. Both of the instruments were designed by consumer psychologists, and focus directly on personality attributes and behavior characteristics directly relevant to purchase and consumption of goods and services. Lastly, the instrument used to measure personality traits was standardized and validated on large numbers of adult consumers, rather than groups such as patients, students, or convicts.

Time Orientation

The individual consumer is faced with a world that is often described in terms of three factors: time, space, and possessions. An extensive literature in the areas of economics and anthropology is devoted to the concepts of property, utility, and the artifacts of man. Geographers and psychologists in the areas of proxemics and kinesics treat human experience in and behavior toward space. The theoretical development and empirical research are less often focused on human existence, experience, and behavior in *time*. The lack of standard measure of time orientation may have inconvenienced, if not inhibited an abundant literature in this area. The requirement for such an instrument was satisfied by the construction of the F-A-S-T Time Orientation Test (Alreck, 1976). The test includes four dimensions of time orientation: *Focus*, *Activity*, *Structure*, and *Tenacity*. The name of the test is an acronym formed from the initials of each scale, and henceforth each will be identified by its initial letter.

Focus. We know from both personal observation of others and from psychological research (Seiden, 1969) that some people tend to direct their consciousness toward the past. They often recall previous experience and relive past occurrences. Others are immersed in the present moment, completely engaged in the here and the now. Still other people can be identified as projecting their thoughts and images to the future and living for what will be. There are also those who see human experience in terms of a span of time, considering the antecedents and consequences, over time, of any present phenomenon. The F scale of the time orientation test was designed to tap this tendency to recollect, to sense, to project, or to spread human consciousness across the time spectrum. It yields a continuous measure of the trait, rather than a typological assignment.

Activity. People differ widely in their perceptions of the supply of time. Some perceive time as passing very slowly, and in effect they have an "over supply" of time, relative to the activities they have available to fill it. At the other end of this spectrum are those who feel that time passes all too quickly, that there are not enough hours in the day. And of course, there are those between these extremes who feel that the supply of time is appropriate to the amount of activity in which they engage themselves. The A scale of the time orientation test measures the degree to which an individual is over- or under-active, relative to the fixed supply of time common to all of us.

Structure. Some people view time as a continuous, smooth, flowing substance; a never ending river flowing from past to present to future. They "take it as it comes," engaging in one activity after another without much regard for the particular time it is or the time it takes to complete a task. Others perceive time as a discrete commodity, neatly packaged in boxes called "noon hours, Wednesday, the 60's," or "the 17th century." Each of these discrete structures is then regarded as appropriate to some particular activity, experience, or phenomenon. Those between the extremes find planning, scheduling, and accounting for time to be acceptable when necessary, but also enjoy a certain amount of spontaneity and freedom from regulated periodicity. The S scale of the time orientation test identifies the location of an individual on this dimension of time experience.

Tenacity. There are those who can engage in an activity almost as though they regarded time as irrelevant so far as external rewards are concerned. These people are able to delay gratification for long periods of time, while pursuing some far-distant goal. Their mirror images are persons who demand rewards almost continuously as their efforts are exerted. Unwilling to wait for the outcomes of their endeavors, they shy away from those tasks or processes that require long delays between sacrifice and receipt of extrinsic reinforcement. Between the extremes are those who find it acceptable to delay gratification for a short period of time, but not for prolonged periods. The T scale provides an indication of each person's willingness to "hang in there" when there is a delay between exertion of effort and receipt of rewards.

Statistical characteristics of F-A-S-T. The test was originally standardized on 930 adult consumers and current norms are based on approximately 2,500 similar subjects. Reliability coefficients for the scales range in the 70's and 80's, while item redundancy and correlations with scale scores are well within acceptable ranges. The scales were validated against spouse observations of time-relevant behavior, job satisfaction with time-related characteristics, and several standardized psychological measures of relevant constructs (Hix, 1976). Focus was reported to be related to Jungian typology (Mann, Siegler and Osmond, 1968) but did not prove to be so when correlated with Myers-Briggs Type Indicator scores (Myers, 1962). Activity was significantly correlated with the time attitude scales of Calabresi and Cohen (1968) while structure was associated with other generalized control constructs (Budner, 1962; Moresko, Schontz and Marrow, 1954). Tenacity proved to be significantly related to the need for achievement measures of Hermans (1970) and Mehrabian (1968). Based on these characteristics of the test and on the constructs measured by the various scales, the test appeared to be an appropriate measure of consumer time orientation. The test had been standardized on consumers, and the traits might be hypothesized to influence consumer behavior.

Consumer Life Style

Before focusing on the relationship between time orientation and more specific aspects of consumer behavior, it seemed appropriate to first determine if time orientation traits influenced the more general patterns of consumer behavior. These patterns could be measured by a library of items relating to consumer Activities, Interests and Opinions (Plummer, 1971; Wells and Tigert, 1971; Plummer, 1974). The emphasis was not on identification of individual activities, interests and opinions that significantly relate to time orientation, but rather on identification of patterns of response that would point to a life style more typical of one type of

time orientation than another. Thus, we hypothesized that each of the four time orientation dimensions would be indicative of a particular life style pattern, as revealed by the activities, interests and opinions.

Methodology

The data to measure the relationships between time orientation and life style patterns were obtained in a field survey of 599 adult consumers in December of 1976. The sample was selected on a convenience basis, with quotas on the distributions of sex, age, and occupational group. As a result, 54 percent of the respondents were male, 64 percent were married, 63 percent were in white collar occupations, and 65 percent indicated they were in the middle class.

The self-administered questionnaire consisted of a cover letter explaining the nature of the project, the F-A-S-T Time Orientation Test, a section containing 406 activity, interest and opinion items, and a final set of questions to obtain the demographic status of the respondents. The questionnaire booklet was twenty pages in length, and was completely self-administered. The field worker contacted the potential respondent, explained the nature of the project, and solicited the person's cooperation. As an inducement to respond, the individual was offered a complete, six-page, computer-generated interpretation of his or her own time orientation. Field workers briefly displayed a copy of the print-out and assured the respondent that most people found it both interesting and informative. If the person decided to participate, he or she was provided with a booklet and given a few days to complete the task. The field worker called back and subsequently picked up the completed questionnaire. To avoid fatigue effects, respondents were encouraged to complete only one or two sections at a sitting, and to set the instrument aside if interrupted or tired.

Each respondent, was, of course, provided with the test interpretations, which proved to be a very potent inducement. The procedure also provided an opportunity for 100 percent validation of the field data collection, since every respondent was contacted a second time. A total of 668 questionnaires were given to potential respondents, with a completion rate of 94 percent. After elimination of those responses which were unacceptable, the final sample size was 599.

The 64 items of the time orientation test were scaled from one to five, with extremes labeled "Exactly like me" and "Not at all like me." After reflecting and con-gregating items, the scoring provided summated scores. Activity, interest and opinion items were scaled from one to six, with extremes labeled "Definitely agree" and "Definitely disagree," or were simply the number of times the individual had engaged in the activity. The relationships between time orientation and AIO items were measured by correlation analysis.

Results

The results of the study are depicted in Tables 1 through 4. Each table is devoted to a single scale of the F-A-S-T Time Orientation Test: *Focus*, *Activity*, *Structure*, or *Tenacity*, respectively. The first column of each table contains the number of AIO items that were significantly correlated with the time orientation test scale at the .01 level and were indicative of a particular aspect or dimension of life style, based on similarity of item content. These life style dimensions are listed in the body of the table in descending order, beginning with the ones that contained the greatest number of items with systematic associations with time orientation.

Focus

The summary of life style dimensions related to time focus is shown in Table 1. Of the total of 406 items, 81 were significantly correlated to the F scale. The AIO items were not highly correlated to the scale, since the largest coefficient was only .32; however, the number of items, together with the fact that the signs of the coefficients were consistently in the appropriate direction, indicates that there is a very clearly discernable relationship between consumer life style and individual time orientation.

Past oriented consumers tend to be the more cautious shoppers. They more often agreed with items such as: *I often worry that something I buy will turn out to be a mistake. I tend to do things pretty much the way my mother did. Grocery shopping for my family means sticking to a strict budget.* On the other hand, the future oriented consumer more often agreed that: *The new styles turn me on. I would like to own the most expensive things.* The past oriented consumer was also less secure, in general, tending more often to agree that: *I dread the future. Things are changing too fast. Modern life is anxiety ridden.* Some items relating to other dimensions were: *I work very hard most of the time.*

TABLE 1
TIME FOCUS AND CONSUMER LIFE STYLE

Items ^a	Past Orientation	Future Orientation
15	Cautious shopper	Consumption oriented
10	Insecure	Secure
9	Negative about work	Positive about work
8	Opinion follower	Opinion leader
6	Homebody	Cosmopolitan
5	Morally rigid	Morally flexible
5	Indifferent to cars	Car conscious
5	Conservative	Innovative
4	Financially pessimistic	Financially optimistic
4	Content	Ambitious
4	Detached	Involved
3	Cautious	Adventurous
3	Stationary	Mobile

^aNumber of AIO items correlated with Focus scale, P less than .01.

I would not work if I did not have to. Future oriented people tended more to agree to the first of these, and past oriented leaned toward the second. So, too, future oriented persons saw themselves as opinion leaders, were generally more cosmopolitan, flexible, and innovative than their past oriented counterparts. They were more ambitious, involved, adventurous, and mobile, while the typical past oriented person was more pessimistic about financial matters and more indifferent to cars, leaning more toward the home.

Activity

The consumer life style dimensions related to activity level are depicted in Table 2. Highly active consumers indicated their involvement on such items as: *I do more things socially than do most people. I am active in community projects.* Low activity people tended to agree: *The old ways are the best ways.* They also more often agreed: *When I retire, I want to sit and take it easy.* Highly active people, on the other hand, had more positive attitudes toward work: *I take pride in my job.* These people were also more family centered, they were

more confident shoppers, and they demonstrated a greater satisfaction with life in general. Low active respondents expressed conservative attitudes toward innovations, indicated more insecurity and uncertainty, and were more often the homebodies. It is also interesting to note that they were more often in agreement with maintenance of sex roles and more indifferent to appearances.

Correlation coefficients between AIO items and the A scale ranged to a high of -.33 for an opinion leadership item: *Nobody cares what I think.* Signs on coefficients were consistently in the predicted direction. In all, there were 89 items from the 406 AIO items that provided a significant correlation with the A scale. Of the four dimensions of time orientation, the activity scale scores were correlated with more items than any of the other three.

TABLE 2
ACTIVITY LEVEL AND CONSUMER LIFE STYLE

Items ^a	Low Activity	High Activity
11	Detached	Involved
10	Opinion follower	Opinion leader
10	Negative about work	Positive about work
10	Self centered	Family centered
9	Apprehensive shopper	Confident shopper
9	Dissatisfied with life	Satisfied with life
7	Conservative	Innovative
5	Insecure	Secure
5	Uncertain	Confident
5	Homebody	Cosmopolitan
5	Past oriented	Future oriented
5	Reluctant to volunteer	Willing to volunteer
4	Sex biased	Unbiased toward sexes
4	Indifferent to appearance	Appearance conscious

^aNumber of AIO items correlated with Activity scale, P less than .01.

Structure

The 10 dimensions of consumer life style defined by 58 AIO items significantly correlated with time structure are shown in Table 3. The dimension containing the largest number of items was related to shopping, with those who were time structured also tending to be careful shoppers and those low on structure leaning toward more casual shopping habits and attitudes. Examples of items positively correlated with structure are: *Before going shopping, I sit down and prepare a complete shopping list. I would not shop in a store that looked dirty, even if the prices were very low. When I get a price off coupon, I save it and use it if I can. I watch the advertisements for announcements of sales. I rely on facts, not emotions when I make decisions.* The AIO items with the highest correlation with the S scale, +.24, was: *My days seem to follow a definite routine, e.g., eating meals at a regular time, etc.* Those high on time structure also tended to agree that: *Visitors often comment on how nice our home or apartment looks. I keep my house very neat and clean.* On the other hand, those who tended not to structure time also agreed: *I don't spend very much time cleaning my house. I find cleaning my house an unpleasant task.*

On the whole, highly structured consumers also tended to be work oriented, principled, and risk averse. They also exhibited more propensity to be financially optimistic and car conscious, while those low on structure

TABLE 3
TIME STRUCTURE AND CONSUMER LIFE STYLE

Items ^a	Low Structure	High Structure
11	Casual shopper	Careful shopper
9	Deemphasis of house-keeping	Emphasis on house-keeping
7	Leisure oriented	Work oriented
5	Flexible	Principled
5	Risk indifferent	Risk aversion
5	Night oriented	Day oriented
5	Follower	Leader
4	Opinion follower	Opinion leader
4	Financially pessimistic	Financially optimistic
3	Car indifferent	Car conscious

^aNumber of AIO items correlated with Structure scale, P less than .01.

were more pessimistic about their financial condition and future and relatively indifferent to cars. Opinion leadership and leadership in general were more typical of highly structured people than those low on structure. Another interesting characteristic of low structure respondents was their preference for activity at night, rather than during the day. Highly structured persons also preferred, to a greater degree, to get up and get going in the morning. They were more often "day" people.

Tenacity

The T scale of the time orientation test provided the weakest systematic relationship to consumer life style. Only 33 of the 406 AIO items were significantly correlated with tenacity scores, and these were grouped into six life style dimensions. Of the six, shopping activities, interests and opinions were the most numerous. The more tenacious the individual, the more likely the person would agree: *I do most of my grocery shopping in one store. When I start to shop for furniture, I usually know exactly what I want. I always buy quality brands.* Those who were less tenacious also less often: *Shopped for household articles. Went shopping for clothes. Returned an unsatisfactory product.*

Other life style differences between low and high tenacious consumers related to housekeeping and attitudes toward the past. Tenacious people also tended to express more compulsive tendencies toward housekeeping: *I think dirty dishes should be washed promptly after each meal. I can usually work for long periods of time around the house without tiring.* Those low on tenacity more often tended to agree: *I dread the future. I dread the unknown. Things are changing too fast.* The tenacious consumer is more assertive: *I am more independent than most people. They agreed less often that: I have never been really outstanding at anything.*

TABLE 4
TENACITY LEVEL AND CONSUMER LIFE STYLE

Items ^a	Low Tenacity	High Tenacity
12	Casual shopper	Structured shopper
6	Unassertive	Assertive
5	Flexible	Conservative
4	Past oriented	Future oriented
3	Self centered	Family centered
3	Casual housekeeper	Compulsive housekeeper

^aNumber of AIO items correlated with Tenacity scale, P less than .01.

Highly tenacious individuals also exhibited more conservatism in: *I am in favor of very strict enforcement of all laws.* More typical of low tenacity was the item: *I like to think I am a bit of a swinger.* The more tenacious, the more family oriented the respondent: *My major hobby is my family.* Even though this dimension of time orientation was less related to the activities, interests and opinions of consumers, the differences in life style are still evident.

Conclusions

The results of the study indicate that individual time orientation is translated into measurable differences in consumer life style. In summary, the following relationships were manifest:

1. The more the individual tends to habitually direct his or her consciousness toward the future, the more consumption oriented the person will be. Such people are also more secure and positive about their work. They tend to be opinion leaders with more cosmopolitan interests. Those who direct their thoughts and fantasies toward the past also tend to be more rigid in their moral outlook, more conservative, cautious, and detached. They exhibit less ambition. They are less mobile than their future oriented counterparts and they are also less car conscious than the future oriented consumer.
2. The highly active person, who perceives time to be in short supply, also tends to be more involved, family centered, and positive about the future. This is the more confident person and that confidence is also manifested in the marketplace. The active person tends to be more secure and satisfied with life than the less active, who tend to be homebodies, less satisfied with their work, and reluctant to volunteer. These less active folks are also more conservative in their views, tend to trust traditional sex roles, and to be more indifferent about their appearance.
3. The greater the consumer's propensity to see time as a continuous, flowing substance, the more likely the person is to display casual shopping patterns, to be leisure oriented, flexible, and indifferent to risk. On the other hand, those who tend to structure time into discrete segments also seem to emphasize house-keeping, regard day for working and night for sleeping, and be more car conscious. Structured consumers value both opinion leadership and leadership in general to a greater degree than those who are less structured.
4. The tenacious individual is more of a structured shopper, assertive and future oriented. The person with less tenacity is also the one who is more likely to be a casual shopper and less compulsive about house-keeping. Tenacity is directly related to family-centered activity and interests and inversely related to flexibility in attitudes and habits.

Obviously, the summary presented above provides more of a caricature of life styles for various time orientations than classifications or typologies. The relationships discovered represent only persistent tendencies and there is a great deal of "overlap" in the life styles of individuals on opposing extremes of any of the time orientation dimensions or scales.

Aside from the more specific findings cited above, the results taken as a whole indicate that "personality" can be related to variations in consumption patterns. This study used a trait-specific personality test that was designed to measure the characteristics of normally functioning people. The test was standardized on the general public, and the "dependent" variables consisted of molar behavior patterns, rather than molecular units,

such as the selection of one brand over another. As expected, the proportions of variance explained were not large in magnitude; however, the associations between time orientation, on the one hand, and life style, on the other, were unmistakably visible. An optimistic interpretation of these results is that personality does influence the behavior of consumers in measurable ways. Further investigations of the role of personality may yet prove fruitful.

References

- Alreck, P.L. *Time Orientation and Behavior*. Unpublished manuscript, San Diego State University, 1976.
- Budner, S. "Intolerance of Ambiguity as a Personality Variable," *Journal of Personality*, 30 (1962), 29-50.
- Calabresi, R., and J. Cohen. "Personality and Time Attitudes," *Journal of Abnormal Psychology*, 73 (1968), 431-439.
- Hermans, H.J.M. "A Questionnaire Measure of Achievement Motivation," *Journal of Applied Psychology*, 54 (1970), 353-363.
- Hix, L.J. *The Personality Correlates to Time Orientation*. Unpublished manuscript, San Diego State University, 1976.
- Kassarjian, H.H. "Personality and Consumer Behavior: A Review," *Journal of Marketing Research*, 8 (1971), 409-418.
- Mann, H., M. Stiegler, and H. Osmond. "The Many Worlds of Time," *Journal of Analytical Psychology*, 13, (1968), 35-56.
- Markin, R.J. *Consumer Behavior*. New York: Macmillan Publishing Co., Inc., 1974.
- Mehrabian, A. "Male and Female Scales of the Tendency to Achieve," *Educational and Psychological Measurement*, 28 (1968) 493-502.
- Moresko, R.M. Rubin, M. Schontz, and W.R. Marrow. "Rigidity of Attitudes Regarding Personal Habits and its Ideological Correlates," *Journal of Abnormal and Social Psychology*, 49 (1954), 89-93.
- Myers, I.B. *The Myers Briggs Type Indicator Manual*. Princeton, New Jersey: Educational Testing Service, 1962.
- Plummer, J.T. "Life Style Patterns and Commercial Bank Credit Card Usage," *Journal of Marketing*, 35 (1971), 35-41.
- _____. "The Concept and Applications of Life Style Segmentation," *Journal of Marketing*, 38 (1974), 33-37.
- Seiden, H.M. *Time Perspectives and Styles of Consciousness: An Exploratory Study*. Unpublished doctoral dissertation, New School for Social Research, 1969.
- Wells, W.D. and D.J. Tigert. "Activities, Interests and Opinions," *Journal of Advertising Research*, 2 (1971), 27-35.

BELIEF SYSTEMS AND THE DIFFERENTIAL ROLE OF THE SELF-CONCEPT

George E. Belch (student), University of California, Los Angeles

Abstract

This study investigates the relationship between belief systems and the differential role of self and ideal self-image in determining purchase intention. The results indicate that the level of conceptual functioning at which an individual operates appears to influence the differential role of the self-concept on purchase intentions. The findings also show that the purchase intention of individuals who are characterized by a high need for cognitive consistency are more influenced by product ownership than those individuals who are more tolerant of cognitive dissonance and inconsistency.

Introduction

Relationships of product image and self-image have been given much attention in consumer research. The idea that consumers prefer products or brands that are congruent with their self-concept has been investigated by many researchers including Birdwell (1968), Grubb and Hupp (1968), Dolich (1969), Ross (1971), and Landon (1974).¹ The basic assumption of these studies has been that the self-concept is related to the manner in which consumers perceive and purchase products. To the extent that the self-concept is of importance to the individual, he will direct his behavior towards maintenance and enhancement of his self-concept.

Evidence supporting a congruence relationship between self-image and purchase behavior has been reported by Birdwell (1968) and Grubb and Hupp (1968). These studies dealt with the consumers matching of product-image with self-image. However, the consumer might be attempting to express and "ideal" self-image through his purchase of a particular product or brand rather than his "actual" self-image. Hamm and Cundiff (1969) found that individuals with a large discrepancy between actual and ideal self-image describe themselves differently in terms of products and have different product perceptions from those individuals with a small discrepancy between the two.

Several studies have examined the congruence between self and ideal self-concept and products or brands. Dolich (1969) found that there was a greater congruence between self-concept and brand most preferred than brand least preferred. He also found that both self- and ideal self-image were almost equally congruent with most preferred brand.

Ross (1971) also investigated the role of self-image and ideal self-image in consumer purchase behavior. He hypothesized that individuals would prefer products which they perceived to be similar to their self-concept. Ross also predicted that in some cases self-image would best describe brand preference. He hypothesized that ideal self-image would be more related to brand preference when the product was more, rather than less, conspicuous to others, and conversely, self-image would be more closely related to brand preference for less conspicuous products.

¹Self-concept, as used in this article, refers to both the "self-image" and "ideal self-image" of the consumer. The latter two terms are viewed as the two components of the self-concept and will generally be given specific reference.

The results of his study strongly supported the prediction concerning subject's preference for brands which were similar to their self-concept. However, little support was found for the predictions regarding the role of self- versus ideal self-image and brand preference. Ross concluded that these results were supportive of a simple congruity relationship between self-concept and brand preference.

Landon (1974) examined the differential role of the self- and ideal self-image and purchase intentions. In this study he was interested in consumer tendencies to match product image with self-image and/or ideal self-image. Landon hypothesized that some individuals would show a dominant influence of self-image/purchase intention correlation over all products, or an actualizing tendency, while some individuals would show a dominant influence of ideal self-image, a perfection tendency. The actualization model views the individual as striving to achieve harmony between his environment and his self-concept by coming to know and accept his self-concept, while the perfection model states that the individual is constantly trying to improve on his self-concept, i.e., achieve his ideal self-concept (Landon, 1974).

The results of Landon's investigation showed that factors related to products and to the individual are important determinants of the relative impact of self-image and ideal self-image on purchase intentions. Over all subjects, purchase intention for some products tends to be more correlated with ideal self-image while for some products purchase intention tends to be more correlated with self-image. Over all products, some subjects were characterized by a higher self-image/purchase intention correlation (actualizers) while some tended to be characterized by a higher ideal self-image/purchase intention correlation (perfectionists). However, neither the actualization nor the perfection hypothesis explained all of the data and Landon concluded that further research is needed to focus on conditions which might explain when each model is likely to be operative.

These studies give some indication that the self-concept is related to product perception. However, it is evident that there is still a great deal of uncertainty concerning the relative roles of the self- and ideal self-image. The findings thus far indicate that self- and ideal self-image are related and may be equally congruent with brand preference, or in some cases brand preference may be more related to one than the other.

The questions of why either self-image or ideal self-image is more related to purchase behavior and why there is a high correlation between the two self-concepts are still unanswered. One possible explanation for the relationship between self-concept and purchase behavior, as well as the degree of real-ideal self-discrepancy, lies in the level of conceptual functioning at which an individual operates. Harvey, Hunt, and Schroder (1961) have developed four belief systems, each representing a different level of conceptual functioning along a concreteness-abstractness dimension. Given the varying characteristics of these four belief systems, each group might behave differently in respect to using either their self-image or ideal self-image as a guide to their purchase behavior for certain products. The purpose of this paper is to examine the possible relationships between belief systems and the differential role of the self-concept in consumer behavior. First, however, it is necessary to examine the general nature

and function of belief systems and how they serve as different systems of motivation.

Belief Systems

A belief system represents a set of predispositions within an individual to perceive, construe, and interpret stimuli or events in a consistent manner. Each individual has certain central or core concepts which guide him in his efforts at making his world consistent. Individuals have conceptual systems which are ways of construing or dimensionalizing relevant aspects of one's world. These systems are characterized by some form of structure or relation among the various parts of the system.

Individuals are believed to develop conceptually through stages, from the undifferentiated and concrete to the more integrated and abstract. The more concrete end of the dimension represents a state of minimal differentiation while abstractness is represented by high differentiation and integration. Individuals will vary according to the degree of differentiation and relativism they attain and will tend to structure their environment according to their own particular conceptual belief system. The variation in differentiation and relativism determines the openness and closedness of the system, its receptivity to deviant elements and its capacity to admit impingements from the outside. Greater abstractness can be viewed as greater openness and greater tolerance of the different and novel (Harvey, 1966).

Harvey, Hunt and Schroder (1961) presented extensive theoretical and empirical bases from which four basic levels of concreteness-abstractness were deduced. The four basic levels were treated as different conceptual systems, each representing a different level of functioning between the hypothetical end points of greater concreteness and abstractness. A brief review of the four belief systems is given before turning to the relationship of conceptual systems to self-concept.

General Characteristics of the Four Systems

System 1. System 1 functioning represents the most concrete mode of dimensionalizing and construing the world. There is a simple cognitive structure with regard to domains of high involvement and this structure is fairly undifferentiated and poorly integrated. Representatives of System 1 show a greater tendency towards extreme and polarized judgments. This system is assumed to have evolved from situations in which individuals have been restricted in the exploration of their environment.

System 1 representatives tend to manifest such characteristics as high absolutism and closedness of beliefs, high evaluativeness, strong adherence to rules, high ethnocentrism, dogmatism and authoritarianism (Harvey, 1966). They also have a tendency to hold beliefs in accordance with clear cut definitions. Representatives of this system have a low tolerance for ambiguity and are negatively aroused by cognitive inconsistencies. This level of functioning is disposed towards system maintenance by excluding potentially conflicting stimuli from the system.

System 2. System 2 functioning is somewhat more differentiated and abstract than System 1. This level of functioning is characterized by negativism and an anti-rule, anti-authority orientation. Members of System 2 tend to be low in self-esteem and high in alienation and cynicism. This system is assumed to result from experiences with an authority who acts omnisciently and is capricious in the control of rewards and punishment. The arbitrary and unpredictable manner in which these sanctions are administered renders the individual unsure of whether courses of action will be reinforced or punished. This results in higher differentiation than System 1,

but also produces a feeling of uncertainty, distrust of authority, and rejection of social norms and guidelines. Representatives of System 2, like those of System 1, have a high need for structure and their tolerance for ambiguity is low. They are also high in dogmatism and absolutism.

System 3. System 3 functioning represents the next to highest level of abstractness treated by Harvey et. al. This system is assumed to result from an environment of over-protection. One or both of the parents act as a buffer between the developing child and his environment, thus preventing him from exploration of his physical and social world. System 3 individuals are fairly proficient in affecting desired outcomes in their environment by having others do things for them.

System 3 individuals have a great need for interaction with others and are concerned with establishing friendships, intragroup consensus, and developing dependency relations in order to avoid the feeling of helplessness and social isolation that would result from their being on their own. They need constant feedback from meaningful people in their environment in order to guide their behavior and attain the acceptance they need. System 3 representatives also develop more positive ties to prevailing social norms than representatives of the other three systems.

System 4. System 4 individuals represent the most abstract and open-minded of the four belief systems. This system is viewed as evolving from an environment in which the child is encouraged to explore his physical and social worlds and to establish and rely on values derived from his own experiences and thought. Members of this system are characterized by a high task orientation, risk taking, creativity, and relativism. They are more tolerant of ambiguity and cognitive inconsistency or conflict and are more flexible and contextual in thought and action.

System 4 individuals have a highly differentiated and integrated cognitive structure. Their set of internal standards are independent of external criteria, in some cases coinciding with social definitions, and in others not.

Belief Systems and Self-Concept

An individual's belief system will have an effect on his perceptions, cognitions, and actions and will affect his psychological processing and behavior. It is likely that the acts of individuals involved in some type of decision process will be affected by their system of beliefs and concepts. Thus, individuals of each system should manifest variations relative to their perceptions, beliefs, and actions.

Considering the varying characteristics of each system, there should be differences in the manner in which self-concept operates for each level of conceptual functioning. Specifically, this investigation is concerned with the differential role of the self-concept (i.e., self- and ideal self-image) in determining purchase intention for the four systems.

The relative importance of self- and ideal self-image to purchase intention and the discrepancy between the two components of self-concept may in part be due to the level of conceptual functioning of the individuals. For example, members of the first two systems are characterized by concrete functioning. They have a difficult time acting "as if" and role playing. These individuals might fail to see any real discrepancy between how they are (self-image) and how they would like to be (ideal self-image). Moreover, any discrepancy between the two self-concepts might tend to create a state of cognitive inconsistency for System 1 and 2 individuals. Thus,

people in these systems would have a small discrepancy between self- and ideal self-image. Members of Systems 3 and 4 should also show congruence between the two self-concepts. However, the discrepancy between self- and ideal self-image for these groups should be larger than for the first two systems.

The purchase intentions of System 3 individuals should be strongly influenced by their ideal self-image. Members of this system have a strong need for affiliation and friendship and will be constantly striving for the attainment of an ideal self-image that will enhance their position among their peers. Members of this system are concerned with social group consensus and peer group influences and may see a certain product and its attributes as a manifestation of the image they hope to project.

System 4 representatives represent the most abstract functioning of the four belief systems. The developing environment of these persons, having been free and open, has offered them an opportunity to explore and possibly discover their self-concept. This group may be more cognizant of their self-image and would no longer feel a need to seek an ideal self-image. Considering the ability of System 4 individuals to be contextual in their thinking, they may be more capable of relating the attributes and image offered by a product to their self-image.

Of primary interest in this study are the relationships between self-image and ideal self-image for the four systems and the role of each in determining purchase intentions for each group. The following research hypotheses were formulated based on the different characteristics of the four conceptual systems.

- H₁: The discrepancy between self- and ideal self-image will be smaller for Systems 1 and 2 than Systems 3 and 4.
- H₂: System 3 representatives will show the greatest influence of ideal self-image on purchase intentions of the four belief systems.
- H₃: System 4 representatives will show the greatest influence of self-image on purchase intentions of the four belief systems.

Also of interest in this study is the manner in which product ownership will affect the four systems. Belch and Landon (1977) found that product ownership affects the self-concept ratings for a product, increasing the likelihood of a product being rated as congruent with both self- and ideal self-image. If a true assessment of the role of self-concept in determining purchase intention is to be made, it is necessary to account for the effect of product ownership.

Considering the variations in the systems, particularly with respect to the need to maintain cognitive consistency, they might be affected differently by product ownership. Previous ownership of a product might influence the self-concept scores for that product or affect the purchase intentions of the individual for that product. As noted by Evans (1968) congruity between ownership and purchase intentions might result from a post-purchase process such as dissonance reduction. Individuals who are less tolerant of cognitive inconsistencies, such as members of System 1 and 2, may be influenced more by product ownership than those who can handle the cognitive conflict. The following hypothesis was formulated to test the effect of product ownership across the four systems:

- H₄: Product ownership will have a greater

influence on purchase intentions for Systems 1 and 2 than for the other two systems.

Method

Measurement of Belief Systems

The instrument used in this study has been developed by Harvey et. al. specifically as a measure of belief systems. The "This I Believe Test" (TIB) is a semi-projective sentence completion test. The TIB requests the individual to indicate his beliefs about a number of socially and personally significant concept referents by completing in two or three sentences the phrase, "This I Believe About _____", the blank being replaced by such referents as religion, friendship, people, or the American way of life among others.

By evaluating subject's responses with regard to several dimensions, including structure and content, subjects can be classified into one of the four belief systems discussed earlier or into an admixture of two or more systems. Test-retest reliabilities for the TIB within one week and after six months have been in the .80's while the interjudge reliability for trained judges in classifying subjects into one of the four systems has been .90 or greater (Harvey, 1966). For this study, subject's responses on the TIB were read and scored by professional readers in the department of Psychology at the University of Colorado. Subjects were classified as being representative of one of the four systems or as admixtures. Only those subjects whose responses classified them as being representative of a pure system were used in this study.

Measurement of Product Perception and Purchase Intentions

This study measured product perception relative to self-image and ideal self-image using a product anchored self-concept measure. This measure yields a congruence relationship between self-image and product-image. Subjects are asked to rate each product on a nine point scale ranging from "very much unlike me" to "very much like me". Ideal self-image for each product was then measured on a scale ranging from "very much unlike I want to be" to very much like I want to be". A similar measure was used by Landon (1971, 1974) for measurement of product perception relative to the self-concept.

Purchase intentions for each product were measured using a five point scale. The first four points on the scale divided a reasonable time range into four intervals. The time dimension for each product varied according to the type of product. The fifth scale point was used to indicate an intention never to buy the product. Product ownership was determined by having the subjects indicate which of the products they owned at the time of the study.

Subjects and Procedure

Subjects for the study were male students enrolled in undergraduate marketing courses at the University of Colorado. The TIB and questionnaire were administered during the regular class periods of each course. After everyone completed the TIB, the questionnaire was administered. Subjects were given twenty minutes to complete the questionnaire. When everyone was finished, the TIBs and questionnaires were collected and the class was given an explanation of the study.

The twelve products used in this study included coffee, after-shave lotion, adult games, imported wine, snow skis, mouthwash, country club membership, sun tan lotion, beer, TV dinners, and deodorant. These products were chosen

from a larger list of products used by Landon (1971). Only those products which had a test-retest correlation of .60 or greater on all three measures were used in this study. Two forms of the questionnaire were given, the first presenting the products in a randomized order, and the second a reverse of the original set.

Four questionnaires were eliminated due to homogeneity of responses by the subjects over all items in the questionnaire. After eliminating these subjects and the 36 admixtures, 124 subjects remained in the sample. The number of subjects in each system was as follows:

System 1 = 53, System 2 = 29, System 3 = 18,
and System 4 = 24.

Results

The primary interest in this study is the variation in the differential role of the self-concept among the four systems. Therefore, the analyses were designed to focus on the individuals comprising each system. The approach taken in some studies (e.g. Ross, 1970) has been to use the averaged scores of the sample over each product. This method is not as useful for revealing group differences since these variations may be confounded or even lost in the aggregation process. The analyses for this investigation were performed by treating an individual's scores on each product as single data points. By focusing on the scores of individual's comprising each system over the twelve products, it is possible to determine whether any differences are, in fact, due to the various systems rather than the products.

In order to test the hypothesis concerning the discrepancy between the two components of self-concept, difference scores between self- and ideal self-image were computed for each subject for each of the twelve products. The discrepancy scores were used as criterion variables in a multivariate Hotelling's T^2 test which is appropriate for testing the significance of the differences between the two self-concept ratings.² This test uses a sample mean difference vector and tests whether the population mean difference vector is significantly different from zero.

This statistic was computed separately for each of the four systems. Table 1 shows the T^2 , F statistics and significance level for each system. The F ratios exceed the .05 level of significance for Systems 1, 2 and 4 indicating that for these three systems the difference between self- and ideal self-image is significant. For System 3, however, the discrepancy between the two components of self-concept is not significant. These results are not supportive of the first hypothesis. In fact, the difference between self- and ideal self-image appears to be greater for Systems 1 and 2 rather than Systems 3 and 4.

TABLE 1
Results of Hotelling T^2 Test for Each System

	System 1 (n=53)	System 2 (n=29)	System 3 (n=18)	System 4 (n=24)
T^2	98.86	78.41	40.88	71.06
F	6.50**	3.97**	1.20	3.09*
Degrees of Freedom	12	12	12	12
	41	17	6	12

* Significant at the .05 level
** Significant at the .01 level

²For a description of Hotelling's T^2 test, see Tatsuoka (1971), p. 81.

To test the hypotheses concerning the differential role of self- and ideal self-image and the effect of product ownership, multiple regressions were run for each system with purchase intention as the dependent variable and self-image, ideal self-image and product ownership as independent variables. Subjects' scores for each product on the four measures were used in the regression equations. Thus, for each system the number of observations is equal to the number of subjects in that group times twelve, the number of products.

By examining the regression coefficients of the three variables it is possible to test the hypotheses concerning the differential role of self-concept and the effect of product ownership on purchase intentions across the different systems. The use of regression analysis in this study is not designed to imply causality but only to determine if there are differences between the systems relative to the association of self- and ideal self-image and purchase intentions. This analysis also allows for a better description of this relationship since it takes into account the effect of product ownership.

Table 2 shows the results of the regression run for each system. The regression coefficients, adjusted R^2 , and the significance level of each variable are presented in this table.³

TABLE 2
REGRESSION COEFFICIENTS FOR SELF-IMAGE, IDEAL SELF-IMAGE AND PRODUCT OWNERSHIP FOR EACH SYSTEM
b values and (Beta)

Variable	System 1	System 2	System 3	System 4
Self-image	.156**	.195**	.185**	.232**
Beta	(.279)	(.338)	(.318)	(.376)
Ideal Self-image	.132**	.102**	.191**	.135**
Beta	(.233)	(.197)	(.340)	(.222)
Product Ownership	.636**	.547**	.027	.319*
Beta	(.212)	(.173)	(.009)	(.106)
Constant	4.62	4.62	4.63	4.75
Adjusted R^2	.367	.363	.380	.364

* Significant at .05 level
** Significant at .01 level

Hypotheses 2 and 3 dealt with the role of self- and ideal self-image in determining purchase intention. More specifically, these hypotheses stated that System 3 representatives will show the greatest relationship between ideal self-image and purchase intention among the four systems, while System 4 will have the strongest relationship between self-image and purchase intention of the four groups. As can be seen in Table 2, the regression coefficient for ideal self-image is the highest for System 3 while the coefficient for self-image is the greatest for System 4. These results are consistent with the hypotheses concerning the differential role of the self-concept across the four systems.

The final hypothesis predicted that product ownership will have a greater influence on purchase intentions for Systems 1 and 2 than the other two groups. Examination of the regression coefficients for product ownership across the four systems (Table 2) supports this hypothesis. Product ownership has large regression weights in Systems 1 and 2 while for Systems 3 and 4 these regression coefficients are smaller in magnitude.

³Both b values (non-standardized) and Betas (standardized) are shown in Table 2. However, the discussion uses b values since this coefficient is required for a subsequent analysis reported below.

The findings discussed thus far are supportive of the hypotheses concerning the differential role of self- and ideal self-image and the effect of product ownership among the four systems. However, to determine whether or not these results are statistically meaningful, all of the data were pooled in an overall regression equation designed to test for the significance of the differences in these specific regression coefficients. The results of this test are shown in Table 3.⁴

TABLE 3
REGRESSION COEFFICIENTS FOR SELF-IMAGE,
IDEAL SELF-IMAGE AND PRODUCT OWNERSHIP
ALL SYSTEMS COMBINED

Variable	b value	Beta	F-ratio
SC	.170	.295	80.62**
SCD4	.064	.094	10.89**
ISC	.124	.226	54.70**
ISCD3	.068	.094	11.96**
OWN	.200	.066	2.83
OWND5	.403	.124	7.55**
Constant	4.64		
Adjusted R ²	.370		

**Significant at the .01 level

In this analysis dummy variables representing specific systems are multiplied with the original variables and included in the regression analysis. The variables SCD4, ISCD3, and OWND5 are cross products between variables SC (self-image), ISC (ideal self-image), and OWN (product ownership) and dummy variables D4, D3, and D5, respectively. Dummy variable D3 represents System 3, D4 System 4, and D5 systems 1 and 2. If the coefficients for the cross product variables are statistically significant, this means that the difference in the value of the regression coefficient for a specific system from the other systems is significant.

For example, self-image (SC) in general has a regression coefficient of .170. However, the coefficient of self-image for System 4 is somewhat greater than the other systems, specifically by .064 which is the value for the coefficient of SCD4. Moreover, this difference is significant beyond the .01 level. The difference in the value of the regression coefficient for ideal self-image between System 3 and the other systems (ISCD3) is .068. This difference is also significant beyond the .01 level.

These results offer statistical support for hypotheses 2 and 3. The influence of self-image on purchase intentions across the four systems is the greatest for System 4 while ideal self-image has the strongest influence for members of System 3.

The differences in the influence of product ownership across the four systems are also tested in this final equation. Overall, the effect of product ownership (shown by OWN in Table 3) is not significant. However, the difference in the influence of ownership for Systems 1 and 2, represented by OWND5 in Table 3, is quite large

⁴The b values are used in this analysis since the differences in the standard deviation of the variable in Table 3 (e.g., SC and SCD4) makes the Betas non-additive.

(.403). This difference is also significant beyond the .01 level. Clearly, product ownership influences purchase intention for Systems 1 and 2 much more than Systems 3 and 4.

Discussion

The results of the present study are interesting in several respects. The findings suggest that predispositional factors related to the level of conceptual functioning of an individual are related to the differential influence of self- and ideal self-image on purchase intentions. It also appears that the impact of product ownership on purchase intentions is mediated by an individual's level of conceptual functioning.

The evidence shows that for Systems 1, 2 and 4 the difference between self- and ideal self-image is significant. However, for System 3 this discrepancy is smaller and not statistically significant. Concreteness of conceptual functioning does not appear to be an explanation for any congruence between the two components of the self-concept.

One possible explanation for the unpredicted large discrepancy between the components of self-concept for Systems 1 and 2 may be that these two groups are void of a well-developed, ideal self-image. Their concrete level of functioning might dispose them towards viewing the "self-image" as the relevant and meaningful component of their self-concept. However, given the nature of the task, which requires subjects to rate each product first with respect to self-image then ideal self-image, representatives of these two systems are confronted with the problem of defining their ideal self-image. If it is not initially well defined, the ideal self-image for these two systems might be more susceptible to a "contrasting effect", wherein the more well-defined self-image serves as an anchor against which the less well-defined ideal self-image is contrasted. This would result in a large discrepancy between self- and ideal self-image for System 1 and 2.

There are, however, differences among the four systems with respect to the influence of either self- or ideal self-image on purchase intentions. Across the four groups, ideal self-image has the greatest impact on purchase intentions for System 3. The conceptual functioning for System 3 representatives predisposes them to be more concerned with such factors as social group consensus and peer group influence. They also have a strong need for affiliation and friendship and tend to develop more positive ties to prevailing social norms. These characteristics might motivate the System 3 individual to attempt to enhance his self-concept through reliance on an image that he feels is consistent with his idealized self. The ideal self-image will be the source of direction for such persons as they search for approval and guidance from others. This suggests that factors such as other-directedness, social desirability of product or brand, and reference group influence might be important determinants of the role of the ideal self-image in guiding purchase behavior.

The results also indicated that System 4 individuals show the strongest influence of self-image on purchase intentions of the four systems. This domination of self-image for System 4 suggests that persons who have had the opportunity to explore and become aware of their self-concept might tend to rely on a relevant self-image as a guide to their purchase intentions. The conceptual functioning of the System 4 individual may be more in line with the actualization model of personality discussed by Landon (1974). As mentioned previously, this model views the individual as striving to achieve harmony between his environment and his self-concept by coming to know and accept his self-concept.

The finding that product ownership had the greatest influence on Systems 1 and 2 is particularly interesting. The possible confounding effects of post-purchase measures is an important issue in self-concept research. The congruity principle of Birdwell (1968) was challenged by Evans (1968) on the grounds that it did not test causality since congruence was measured after the purchase decision was made. Evans argued that post-purchase data does not necessarily reflect purchase motivations due to self-image/product-image congruity. After purchase congruity or indication of future purchase intentions might result from a process such as dissonance reduction.

The concreteness of conceptual functioning of System 1 and 2, which disposes them toward a lower tolerance for dissonance, may motivate them to give more consideration to the fact that they presently own the product under consideration. The greater influence of product ownership for these two systems might reflect a way of maintaining cognitive consistency between past and future purchase behavior. Other empirical evidence (Harvey, 1965; Ware and Harvey, 1967; Harvey and Ware, 1967) has shown that dissonance and consistency theories have greater validity for concretely functioning than abstractly functioning individuals. These results appear to be consistent with this evidence.

The purpose of this paper was to determine if belief systems can offer insight into the differential role of the self-concept in determining purchase intentions. The present study has shown that factors related to an individual's level of conceptual functioning are correlates of the differential influence of self- and ideal self-image on purchase intentions. The conceptual system of the individual appears to mediate the relationship between self-concept and purchase behavior.

This study also indicates that product ownership may affect the self-concept/purchase behavior relationship. The significant influence of ownership on the first two systems suggests that further research might examine the effect of dissonance or other consistency theories on the self-product-image congruity of consumers.

Future research on the role of the self-concept may find it fruitful to concentrate on individual differences such as those suggested in this study and their relationship to the differential role of the self-concept. Finally, this study has shown that belief systems can provide some interesting insights into the behavior of the consumer. The conceptual systems approach of Harvey et. al. may have potential for other areas of research in consumer behavior.

References

- George E. Belch and E. Laird Landon, Jr., "discriminant Validity of a Product-Anchored Self-Concept Measure," Journal of Marketing Research, 14 (May 1977) 252-6.
- Al. E. Birdwell, "A Study of the Influence of Image and Congruence on Consumer Choice," Journal of Business, 41 (January 1968), 76-88.
- Ira J. Dolich, "Congruence Relationships Between Self Images and Product Brands," Journal of Marketing Research, 6 (February 1969), 80-4.
- Franklin B. Evans, "Automobiles and Self-Imagery: Comment," Journal of Business, 41 (October 1968), 445-59.
- Edward L. Grubb and Gregg Hupp, "Perception of Self, Generalized Stereotypes and Brand Selection," Journal of Marketing Research, 5 (February 1968), 125-7.
- O.J. Harvey, "Some Situational and Cognitive Determinants of Dissonance Resolution," Journal of Personality and Social Psychology, 1965, 1, 349-55.
- O.J. Harvey, "System structure, flexibility and creativity," in O.J. Harvey (Ed.), Experience, Structure and Adaptability, New York: Springer, 1966.
- O.J. Harvey, D.E. Hunt, and H.M. Schroeder, Conceptual Systems and Personality Organization, New York: Wiley, 1961.
- O.J. Harvey and R. Ware, "Personality Differences in Dissonance Resolution," Journal of Personality and Social Psychology, 1967, 5, 38-44.
- E. Laird Landon, Jr., "Need for Achievement, Self-Concept, and Product Perception," unpublished doctoral dissertation, University of California, Los Angeles, 1971.
- E. Laird Landon, Jr., "The Differential Role of Self-Concept and Ideal Self-Concept in Consumer Purchase Behavior," Journal of Consumer Research, 1 (Sept. 1974).
- Ivan Ross, "Self-Concept and Brand Preference," Journal of Business, 44, (Jan. 1971).
- Maurice M. Tatsuoka, Multivariate Analysis: Techniques for Educational and Psychological Research, John Wiley and Sons, Inc. 1971.
- R. Ware and O.J. Harvey, "A Cognitive Determinant of Impression Formation," Journal of Personality and Social Psychology, 1967, 5, 38-44.

MASCULINITY AND FEMININITY FACTORS
IN PRODUCT PERCEPTION AND SELF IMAGE¹

James W. Gentry, Mildred Doering, Terrence V. O'Brien
Kansas State University

Abstract

Sex roles are changing in the United States and it has been hypothesized that such changes will be evidenced in new patterns of consumer behavior. This study investigates self image regarding masculinity and femininity, sex identifications of representative products and life style elements, and usage rates of products.

Introduction

Tucker (1976, p. 353) has pointed out the impact of changing sex roles on marketing: "During most of the rest of the century marketers will increasingly miss the center of their markets because they will not understand them. And the change in the relations of the sexes will be the primary cause." This trend was also noted by Stuteville (1971), Hawkins and Coney (1976), and Dickens and Chappell (1977). Stuteville (1971) discusses the cases of two products, one of which (cigarettes) was once used almost solely by males but is now used as frequently by females (Gentry and Doering, 1977), and the other (hair spray) was once used almost solely by females but is now used almost as frequently by males (Gentry and Doering, 1977).

That sex roles are changing in our society should be sufficient for the marketer to become less interested in the male-female dichotomy and more interested in the level of masculinity or femininity. As Sechrest (1976) states, "That a respondent is male is in itself of little psychological impact. What is of consequence are the patterns of attitudes or abilities or problems or interests that are presumed to go along with being male." Aiken (1963) was one of the first to relate masculinity-femininity to purchasing behavior. He found that more feminine women were more likely to belong to the "Decoration," "Interest," and "Conformity" clusters of female dress buyers. Masculinity-femininity was not related to membership in the "Comfort" and "Economy" clusters. Vitz and Johnston (1965) found that, within each sex, a person's masculinity (femininity) is positively related to the masculine (feminine) image of the cigarette smoked. Fry (1971) found that more feminine men were more likely to smoke cigarettes that were identified as being less masculine. Morris and Cundiff (1971) found that masculinity-femininity was not related to the rating of hair spray by males; however, a strong interaction between manifest anxiety and masculinity-femininity was found.

Thus Aiken (1973), Vitz and Johnston (1965), and Fry (1971) found that their subjects' behavior was consistent with their masculine-feminine self image. On the other hand, Morris and Cundiff (1971) found that males with a relatively high feminine identification and a high anxiety level expressed strongly unfavorable attitudes toward the use of hair spray, which was viewed as being relatively feminine.

Stuteville (1971) hypothesized that the consistency of purchase behavior with one's masculine-feminine self image may vary across sexes. He stated that it is easier for a male-oriented product to attract female buyers than the reverse situation, since our culture labels the boy who acts like a girl a "sissy," but the girl who acts like a boy is called a "tomboy." The latter is much easier for a girl to accept than the former is for a boy.

The purpose of this study is to explore male and female perceptions of a wide variety of products and leisure activities and to explore whether one's masculinity-femininity self image within each sex is related to the individual's product perceptions. Given the convergence of sex roles in our society, we hypothesize that purchase behavior will be consistent with the masculine-feminine self image for 1) sexually neutral items and for 2) items associated with the opposite sex, but that inconsistency will occur for 3) items identified with the respondent's own sex due to the less threatening nature of those items.

Methodology

The population chosen for study was college students. One reason was, of course, convenience; more important was the belief that changing sex roles in our society (and corresponding changes in purchasing behavior) might be more evident in college students than in their elders.

The selection of specific leisure activities and products to be evaluated was done through a three-stage process. First, a list was compiled of activities and products that were thought to be related to the life-styles of the college student subjects. Second, a pretest was run concerning attitudes toward the stimuli and frequency of usage of them. A second pretest was run to establish the perceived masculinity-femininity of the stimuli. Stimuli were deleted if they were used very infrequently by most of the respondents. The third stage consisted of a final pretest of the instrument after the deletions were made. Then, additional stimuli were deleted because of similarity to other stimuli and because of the need to shorten the instrument. The specific leisure activities and products used in the study are listed in Tables 1 and 2.

Two hundred college students at Kansas State University (100 males and 100 females) were recruited through an advertisement in the student newspaper. While there was no attempt to randomly select subjects, the procedure did result in a representative mix of curricula. Further, the subjects' scores on a standardized personality instrument (the California Psychological Inventory Fe-scale) were nearly identical with results published for college students in Gough (1975). The mean and standard deviation for the females (23.22 and 2.85, respectively) and males (16.35 and 3.82, respectively) in our study closely approximate the CPI norms (23.16, 3.27; and 16.65, 3.73; respectively). Consequently, there is some evidence that the sample was fairly representative of college students.

Each subject was paid four dollars for completing the series of questionnaires dealing with their ratings of different consumer products and leisure-time activities.

¹The authors would like to thank Rene Klassen for his help in the data collection, Gail Holtman for her help in the data analysis, and the Bureau of General Research, Kansas State University, for financial support of the study.

The subjects also completed two masculinity-femininity scales, the CPI Fe-scale (Gough, 1975) and the Personal Attributes Scale (PAQ: Spence, et al., 1974). The last section of the instrument included demographic questions, usage measures of various media (not analyzed here), and ratings of the products and activities as to their perceived masculinity or femininity. Each subject was told to complete all forms, and each of the 56 sheets of information per subject was checked by a monitor before the subject was paid.

TABLE 1
USE OF PRODUCTS BY SEX OF RESPONDENT AND BY PERCEPTIONS OF PRODUCTS AS MASCULINE OR FEMININE

Product	Use of product, by sex		Use of product, by perception	
	Male	Female	Mas.	Fem.
Beer	4.0 ^a	3.5 ^b	3.4	3.6
Boots	4.8	3.5 ^b	4.2	4.2
Cigarettes	.9	.4	.5	.3
Cologne	4.0	5.2 ^b	4.9	4.4 ^c
Deodorant	4.6	5.3 ^b	4.9	5.0
Hair spray	1.3	1.5	1.6	1.7
Jeans	6.0	6.3	6.2	6.1
Mouthwash	2.5	2.5	2.4	2.6
Razor blades	3.8	3.3	3.5	3.7
Sandals	2.7	6.5 ^b	4.8	4.4
Shampoo	4.2	4.5	4.2	4.4
Soap	5.1	5.6 ^b	5.4	5.4
Tennis shoes	4.6	4.0 ^b	4.2	4.6

^a Entries are group means and are scaled differently for different variables; these include seven point bipolar adjective scales, rate of consumption of the product, and others.

^b The difference between male and female use is significant at the .10 level or better.

^c The difference in usage between masculine and feminine product perception is significant at the .10 level or better.

Discussion of Scales

Gough's (1975) CPI Fe-scale consists of 38 true-false items. This procedure views "masculinity-femininity" as a continuum, and a low score indicates masculinity while a high score indicates femininity. This scale was chosen because of its long tradition, being published first in 1952 (Gough, 1952), and because it has been commonly used in marketing studies (Aiken, 1963; Fry, 1971; Morris and Cundiff, 1971; Vitz and Johnson, 1965).

The PAQ views masculinity and femininity as two distinct dimensions, which permits individuals to be identified as being "androgynous," or masculine and feminine. The androgynous classification is applied to those who are both independent and tender, assertive and yielding, tough but sweet. The first androgyny scale was developed by Bem (1974), but we preferred the PAQ because of the way "androgyny" is defined operationally. The Bem Sex Role Inventory (BSRI) identifies someone as "androgynous" if his/her score on the masculine and feminine scales are comparable. Thus the BSRI identifies respondents as masculine, feminine, or androgynous. The PAQ is

more restrictive in the identification of androgynous individuals; only those who score high on both the masculine and feminine dimensions are identified as being androgynous. Those ranking low on both dimensions are grouped in a fourth category, labeled low-masculine, low-feminine. The only marketing study that we have seen that used the concept of androgyny was that by Tucker (1976), and he also used the PAQ rather than the BSRI.

TABLE 2
PARTICIPATION IN LEISURE ACTIVITIES, BY SEX OF RESPONDENT AND BY PERCEPTIONS OF LEISURE ACTIVITIES AS MASCULINE OR FEMININE

Activity	Participation in activity, by sex		Participation in activity, by perception	
	Male	Female	Mas.	Fem.
Ballet	.7 ^a	2.1 ^b	1.2	.4
Basketball	10.5	9.7	11.9	7.9 ^c
Bicycling	30.5	16.3	22.6	24.3
Car races	2.6	1.5 ^b	1.9	2.4
Fishing	6.4	2.0 ^b	3.6	3.4
Hunting	4.8	.4 ^b	2.2	1.5
Ice skating	2.8	1.0	1.4	.3
Knitting	.4	15.3 ^b	6.0	7.9
Swimming	20.6	24.2	22.2	15.2
Movies, G rated	2.1	3.0 ^b	2.8	2.2 ^c
Movies, PG rated	3.5	4.6 ^b	3.9	3.9
Movies, R rated	2.7	2.7	3.8	4.4 ^c
Movies, X rated	2.4	1.1 ^b	2.1	1.7

^a Entries are group means and are scaled differently for different variables; these include seven point bipolar adjective scales, rate of participation in the activity, and others.

^b The difference between male and female participation is significant at the .10 level or better.

^c The difference in usage between masculine and feminine product perception is significant at the .10 level or better.

The attitude measures consisted, in part, of the four seven-point, bipolar scales used in the Morris and Cundiff study (1971): valuable-worthless, sociable-unsociable, nice-awful, and useless-useful. In addition, an overall product or activity attitude was obtained using a seven-point bipolar scale with very unfavorable attitude and very favorable attitude as the end points. Numerical frequencies were used where applicable, usually with the leisure activities. Most of the product usage questions required the subject to select a response such as "once a month" or "once a day."

Analysis

The respondents were grouped in different masculinity-femininity categories based upon their scores on the two scales. For the CPI Fe-scale, boundaries were established at approximately the 33rd and 66th percentiles. The groups are labeled Masculine (60 males, 2 females), Neither (33 males, 39 females) and Feminine

(7 males, 59 females). For the PAQ, the subject population was split at the median on both the male-valued and female-valued scales (59 and 50, respectively). The four resulting PAQ groups are low masculine and low feminine (32 females, 25 males), low masculine and high feminine (32 females, 13 males), high masculine and low feminine (10 females, 34 males), and high masculine, high feminine (26 females, 28 males).

Each respondent's ratings of the masculinity-femininity of the leisure activities and of the products were normalized to remove differential response-scale tendencies before aggregating over subject. Then each respondent was categorized as viewing the product or activity as being more feminine or more masculine than average. The respondents' usage of the stimuli were then related to their sex, their masculinity-femininity, and how masculine or feminine they viewed the stimuli through a series of analyses of variance. The particular approach used was the Least Squares program written by Kemp (1976), which is designed to handle unequal cell sizes.

The categorization of the groups as being consistent or inconsistent with their self images was based on the mean usages for the different items. The differences in the cells (for example, the difference between the mean usage for those males who viewed the stimulus as being more masculine and the mean usage for the males who viewed the stimulus as being more feminine) were tested using t-statistics.

Results

Product stimuli are shown on the left side of Table 1. Note that only one product (cologne) has a statistically significant difference in the usage rate between masculine-feminine product perceptions. Those who view cologne as being masculine use the product more than those who view it as being feminine. One reason for the lack of statistical significance in usage by those with different masculine-feminine product perceptions may be that the products are essentially unisexual in nature. Another reason is the high variance in the usage rates of the different stimuli. Much greater differences exist, however, in product use by sex of respondent, with seven out of thirteen items showing marked differences. Beer, boots, and tennis tend to be used more by males, while cigarettes, deodorant, sandals, and soap are used more by females.

Similar findings (Table 2) resulted for leisure activities, in that only a few were engaged in differentially by those with different masculine-feminine product perceptions but most were, in fact, engaged in differentially by sex. Reported participation was higher for women in watching ballet, knitting, and attending G-rated and PG-rated movies; it was higher for men for car races, fishing, hunting, and attending X-rated movies. Neutral activities were basketball games, bicycling, ice skating, swimming, and attending R-rated movies. However, for only three items were usage rates different for those with different masculine-feminine product perceptions. Those who viewed basketball games and G-rated movies as being more masculine attended these activities more often. More frequent attendees of R-rated movies tended to view them as being more feminine. It may be that sex role guidelines are not relevant or usually salient for many activities or products, or there may be suppression of their perceived impact.

Table 3 shows product usage frequencies broken down by sex and by whether the product was viewed as being masculine or feminine. The interactions between sex and the perception of the product were significant for boots, shampoo, and soap. For boots and shampoo, the

significant interaction indicated that both sexes were consistent with their self image; that is, males were heavier users if they viewed the product as being more masculine, while females were heavier users if they viewed the product as being more feminine. On the other hand, the interaction for bar soap finds that both sexes are inconsistent with their self-image.

TABLE 3
PRODUCT USAGE FREQUENCIES, BY SEX OF RESPONDENT
AND BY MASCULINITY-FEMININITY PERCEPTION

Product	Perception of Product	Sex of respondent	
		Male	Female
Beer	Masculine	3.3 ^a	3.5
	Feminine	3.7	3.6
Boots ^b	Masculine	5.1	3.2
	Feminine	4.5	3.8
Cigarettes	Masculine	1.3	1.1
	Feminine	1.2	1.2
Cologne	Masculine	4.4 ^c	5.3
	Feminine	3.6	5.1
Deodorant	Masculine	4.6	5.2
	Feminine	4.6	5.3
Hair Spray	Masculine	1.3	1.9
	Feminine	1.5	2.0
Jeans	Masculine	5.9	6.5
	Feminine	6.1	6.1
Mouthwash	Masculine	2.6	2.1
	Feminine	2.4	2.9
Razor Blades	Masculine	3.5	3.5
	Feminine	3.2	4.0
Sandals	Masculine	3.0	6.6
	Feminine	2.5	6.3
Shampoo ^b	Masculine	4.3	4.1 ^c
	Feminine	4.1	4.8
Soap ^b	Masculine	5.0	5.8
	Feminine	5.3	5.5
Tennis Shoes	Masculine	4.3	4.1
	Feminine	4.6	4.6

^aEntries are group means and are scaled differently for different variables; these include seven point bipolar adjective scales, rate of consumption of the product, and others.

^bSignificant at the .10 level or better by analysis of variance.

^cFor this sex of respondent, the difference in usage between masculine and feminine product perceptions is significant at the .10 level or better, by t-test.

The interaction term is most likely to be significant if both sexes are very consistent or if they are both inconsistent. If one sex is consistent while the other one is not, the interaction term is less likely to be significant. To investigate these situations more closely, the differences in the cell means were tested using t-tests. Males were found to be significantly consistent in their use of cologne, and females were found to be significantly consistent in their use of shampoo.

Table 4 shows the usage frequencies of the leisure activities, broken down by sex and by the masculinity-femininity perception of the activity. Attendance of G-rated movies yielded the only significant inter-

action; females were more likely to attend G-rated movies if they viewed them as being more masculine; males attendance did not depend on their sex-role perceptions. On the other hand, females were significantly more likely to attend R-rated movies if they viewed them as being more feminine.

TABLE 4
LEISURE ACTIVITY USAGE, BY SEX OF RESPONDENT
AND BY MASCULINITY-FEMININITY PERCEPTION

Product	Perception of activity	Sex of respondent	
		Male	Female
Ballet	Masculine	.6 ^a	1.8
	Feminine	.3	.6
Basketball	Masculine	12.1	8.7
	Feminine	8.9	6.3
Bicycling	Masculine	23.3	21.9
	Feminine	37.7	10.9
Car Races	Masculine	1.9	2.0
	Feminine	1.9	2.9
Fishing	Masculine	4.8	2.4
	Feminine	3.9	3.0
Hunting	Masculine	3.2	1.2
	Feminine	2.4	.7
Ice Skating	Masculine	2.4	1.0
	Feminine	.4	.5
Knitting	Masculine	.03	11.7
	Feminine	.00	14.8
Swimming	Masculine	22.6	21.8
	Feminine	10.9	19.5
Movies, G Rated ^b	Masculine	2.6	2.9
	Feminine	2.6	1.7
Movies, PG Rated	Masculine	3.7	4.2
	Feminine	3.6	4.1
Movies, R Rated	Masculine	3.7	4.0 ^c
	Feminine	3.9	5.0
Movies, X Rated	Masculine	2.6	1.6
	Feminine	2.1	1.4

^a Entries are group means and are scaled differently for different variables; these include seven point bipolar adjective scales, rate of participation in the activity, and others.

^b Significant at the .10 level or better of analysis of variance.

^c For this sex of respondent, the difference in usage between masculine and feminine activity perceptions is significant at the .10 level or better, by t-test.

Usage patterns were investigated as to self image differences by a standardized personality inventory. Table 5, results are shown from the California Psychological Inventory's Fe-scale for our product set and in Table 6 for the leisure activities. Again, usage pattern differences are not striking.

The ANOVA results indicate that the respondents were consistent with their self-image in the usage of bar soap and attendance of G-rated movies; for example, masculine respondents used bar soap more if they perceived it as being more masculine and feminine respondents used bar soap more if they perceived it as being more feminine. On the other hand, the respondents were

inconsistent with their self-image in the usage of mouthwash. When the different CPI masculinity-femininity categories were analyzed by themselves, the statistically significant findings were that the masculine respondents attend G-rated movies more if they view them to be more masculine (consistent) while feminine respondents attend basketball games more if they view them to be more masculine (inconsistent).

TABLE 5
PRODUCT USAGE, BY RESPONDENT MASCULINITY-FEMININITY
PERSONALITY SCORE AND PRODUCT PERCEPTION

Product	Perception of Product	Personality Orientation of respondent		
		Masculine	Neither	Feminine
Beer	Masculine	3.9 ^a	3.7	2.6
	Feminine	4.2	3.6	3.1
Boots	Masculine	5.6	3.6	3.6
	Feminine	4.7	3.9	3.9
Cigarettes	Masculine	.3	.8	.5
	Feminine	.4	.3	.3
Cologne	Masculine	4.9	4.8	5.0
	Feminine	4.3	4.4	4.5
Deodorant	Masculine	4.9	4.8	4.9
	Feminine	5.1	4.9	5.0
Hair Spray ^b	Masculine	2.2	1.1	1.5
	Feminine	2.1	1.8	1.4
Jeans	Masculine	6.7	5.7	6.3
	Feminine	6.2	5.9	6.2
Mouthwash ^b	Masculine	1.6	3.0	2.6
	Feminine	3.2	2.6	2.1
Razor Blades	Masculine	4.4	3.2	3.0
	Feminine	4.5	3.5	3.0
Sandals	Masculine	5.2	4.7	4.5
	Feminine	4.6	4.5	4.1
Shampoo	Masculine	4.4	4.2	4.2
	Feminine	5.0	4.3	4.0
Soap ^b	Masculine	5.8	5.1	5.2
	Feminine	5.4	5.3	5.4
Tennis Shoes	Masculine	4.7	4.4	3.4
	Feminine	5.1	4.5	4.1

^a Entries are group means and are scaled differently for different variable; these include seven point bipolar adjective scales, rate of consumption of the product, and others.

^b Significant at the .10 level or better by analysis of variance.

Based on use and participation rates reported in Tables 1 and 2, products and activities were classified as masculine, feminine, or neutral, as shown in Table 7. In addition, two redefinitions of previous variables were also included to better measure their effects. In Table 7, quantity of beer consumed per time period is added to the frequency of consumption measure used thus far; also, a dichotomous smoking variable is added to the previous quantity measure. In addition, the results for two of the four PAQ categories are added: the group that scored low on the masculinity scale and high on the femininity scale (the feminine group) and the group that scored high on the masculinity scale and low on the femininity scale (the masculine group). A summary of the other two groups (low masculine-low feminine and high masculine-high feminine) is not provided as their self-images can not be classified as either masculine

or feminine.

TABLE 6

LEISURE ACTIVITY USAGE, BY RESPONDENT MASCULINITY-FEMININITY PERSONALITY SCORE AND PRODUCT PERCEPTION

Product	Perception of Product	Personality Orientation of respondent		
		Masculine	Neither	Feminine
Ballet	Masculine	.2 ^a	.7	2.7
	Feminine	.3	1.1	.1
Basketball	Masculine	7.8	10.1	13.1 ^c
	Feminine	6.6	9.1	7.2
Bicycling	Masculine	21.7	30.6	15.7
	Feminine	12.0	40.3	20.7
Car Races	Masculine	3.6	1.1	1.1
	Feminine	3.8	2.3	1.1
Fishing	Masculine	6.8	2.1	1.9
	Feminine	5.8	2.8	1.7
Hunting	Masculine	5.6	.7	.2
	Feminine	3.2	1.1	.2
Ice Skating	Masculine	3.5	.8	.1
	Feminine	.1	.5	1.3
Knitting	Masculine	.04	8.2	9.5
	Feminine	1.1	.9	21.4
Swimming	Masculine	18.9	23.3	24.3
	Feminine	16.7	16.2	12.9
Movies, G Rated ^b	Masculine	2.6 ^c	2.5	3.3
	Feminine	.7	2.1	3.8
Movies, PG Rated	Masculine	3.5	3.9	4.5
	Feminine	3.6	3.8	4.2
Movies, R Rated	Masculine	4.5	4.0	3.0
	Feminine	5.1	4.4	3.8
Movies, X Rated	Masculine	3.0	1.4	1.9
	Feminine	2.2	1.6	1.5

^a Entries are group means and are scaled differently for different variable; these include seven point bipolar adjective scales, rate of participation in the activity, and others.

^b Significant at the .10 level or better by analysis of variance.

^c For this personality orientation of respondent, the difference in usage between masculine and feminine activity perceptions is significant at the .10 level or better, by t-test.

Entries in Table 7 refer to the consistency of the masculinity-femininity product or activity perception with sex of the respondent and with the respondent's masculinity-femininity personality disposition as measured by two different personality inventories. Again, our hypotheses were that consistency would exist for neutral items and when the product is associated with the opposite sex, but that inconsistency would occur for products and activities identified with the respondent's own sex--the latter because of a pull toward the center as sex roles loosen up and because of the less threatening nature of these items than of those identified with the opposite sex.

Findings for males in Table 7 tend to support expectations for sex of respondent and for the PAQ measure, but not for the CPI. For example, males were more likely to be inconsistent for masculine items (3 out of

7) than for feminine items (1 out of 6) and females were more likely to be inconsistent for feminine items (6 out of 8) than for masculine items (3 out of 8). The feminine groups as categorized by both the CPI and PAQ scales tend to support the hypothesis; they were inconsistent for more feminine items (4 out of 8 for the CPI and 6 out of 8 for the PAQ) than for masculine items (2 out of 7 for the CPI and 3 out of 8 for the PAQ). The masculine group for the PAQ was inconsistent for 5 of the 8 masculine items, but only for 1 of the 8 feminine items. On the other hand, the CPI masculine group was more inconsistent for feminine items (4 of 8) than for masculine items (3 of 7).

TABLE 7

CONSISTENCY OF MASCULINITY-FEMININITY PRODUCT AND ACTIVITY PERCEPTION WITH SEX OF RESPONDENT AND WITH PERSONALITY ORIENTATION OF RESPONDENT, BY DOMINANT USER TYPE

Product or activity	Personality disposition of respondent					
	Sex of respondent		CPI scale		PAQ scale	
	M	F	M	F	M	F
<u>Masculine items^a</u>						
Beer - frequency	- ^b	+	-	+	-	-
- quantity	-	-	0	+	-	-
Boots	+	+	+	+	-	+
Tennis shoes	-	+	-	+	-	+
Car races	0	+	-	0	-	+
Fishing	+	+	+	-	+	-
Hunting	+	-	+	0	+	+
Movies, X rated	+	-	+	-	+	+
<u>Neutral items</u>						
Hair spray	-	+	+	-	+	+
Jeans	-	-	+	-	+	-
Mouthwash	+	+	-	-	-	-
Razor blades	+	+	0	0	-	+
Shampoo	+	+	-	-	-	+
Cigarettes-quantity	+	+	-	-	+	-
-smoke or not	+	+	-	-	+	-
Basketball	+	-	+	-	+	-
Bicycling	-	-	+	+	-	-
Ice skating	+	-	+	+	-	+
Swimming	+	-	+	-	+	+
Movies, R rated	-	+	-	+	+	+
<u>Feminine items</u>						
Cologne	+	-	+	-	+	-
Deodorant	0	+	-	+	+	-
Sandals	+	-	+	-	-	-
Soap	-	-	+	+	+	+
Ballet	+	-	-	-	+	-
Knitting	+	+	-	+	+	+
Movies, G rated	0	-	+	+	+	-
Movies, PG rated	+	-	-	-	+	-

^a Designation of items as masculine, neutral, or feminine is based on product use and activity participation results from Tables 1 and 2.

^b "+" means that heavier users of this type (male, masculine, etc.) view the item as being consistent with their orientation, "-" means that heavier users of this type view the item as being inconsistent with their orientation, and "0" means that there is no difference in usage between those who view items as being masculine and those who view it as being feminine.

In general, the expectations do have some support when the items categorized as masculine or feminine are considered. However, we hypothesized that the respondents would be consistent with their self-images in the usage of neutral items; no such general pattern can be observed in Table 7. The only group for which this appears to have support is for males; the other groups are apparently just as likely or even more likely to be inconsistent with their self-concepts in their usage of the neutral items. Aggregating over items, it seems that males tend to be more consistent with their self-image (17 of 25 cases) than do females (14 of 28 cases). We would have to conclude that the evidence is perhaps suggestive as to general trends but well short of persuasive, especially when the scarcity of statistically significant results is considered.

Finally, a comparison can be derived from Table 7 for the personality measures, which shows little correspondence between the two. For example, there are ten instances of consistency registering on both the CPI and PAQ scales, five of inconsistency, but eleven mismatches (plus on one and minus on the other) occur. It is not uncommon to find both the masculine and feminine groups from one scale to be classified as being consistent, while the counterparts from the other scale are both classified as being inconsistent. The two scales are measuring somewhat different phenomena and the differences are evident here.

Conclusion

This study looked at a demographic variable (sex) and two measures of masculinity-femininity self image, the California Psychological Inventory Fe-scale and the Personal Attributes Questionnaire, as these related to representative marketing stimuli. It was hypothesized that consumers will use products with neutral orientations or orientations opposite from theirs (for example, feminine products for males) as long as they perceive the products to be consistent with their self-image. It was further hypothesized that heavier users of products with similar orientations to their own (for example, feminine products and females) will perceive the product to be inconsistent with their self image. The results provide very weak support for the directions predicted by the hypotheses.

Implications

It may well be that the marketing stimuli and the leisure activities selected for this research were too "unisexual" in nature, the result being not very striking findings on sex role differences. More extreme stimuli might have more strongly supported our hypotheses. We would expect, however, that a trend toward less stereotypic sex roles still could be identified, although the ideal research design for this purpose would consist of a demonstrably representative set of stimuli along with a longitudinal measurement structure. Future work in this area will hopefully include some of these features.

Our findings do lend weak support to the convergence of sex roles in consumer behavior. The lack of signifi-

cant differences in the usages of products that were formerly sex-stereotyped (cigarettes, jeans, hair spray, and shampoo) lends support to the claims of developing unisexuality of many products. Also, the weak support for our hypothesis that consumer behavior will be inconsistent with the masculine-feminine self image when the stimulus is associated with the same sex and consistent when associated with the opposite sex lends support for the convergence of sex roles. However, as noted in Gentry and Doering (1977), the male-female dichotomy still explains more of the variability in consumer behavior than does masculinity-femininity.

References

- Lewis R. Aiken, Jr., "The Relationships of Dress to Selected Measures of Personality in Undergraduate Women," Journal of Social Psychology, 59 (1963), 119-128.
- Frank M. Bass, Edgar A. Pessemier, and Douglas J. Tigert, "A Taxonomy of Magazine Readership Applied to Problems in Marketing Strategy and Media Selection," Journal of Business, 42 (July 1969), 357-363.
- Sarah L. Bem, "The Measurement of Psychological Androgyny," Journal of Consulting and Clinical Psychology, 42 (1974), 155-162.
- Jackie Dickens and Brian Chappell, "Food for Freud? A Study of the Sexual Polarisation of Food and Food Products," Journal of the Market Research Society, 19 (No. 2, April 1977).
- Joseph N. Fry, "Personality Variables and Cigarette Brand Choice," Journal of Marketing Research, 8 (August 1971), 298-304.
- James W. Gentry and Mildred Doering, "Masculinity-Femininity Related to Consumer Choice." Proceedings, Fall Conference, American Marketing Association, 1977, Forthcoming.
- Harrison G. Gough, "Identifying Psychological Femininity," Educational and Psychological Measurement, 12 (1952), 427-439.
- Harrison G. Gough, Manual for the California Psychological Inventory, Palo Alto, Cal.: Consulting Psychologists Press, Inc., 1975.
- Del I. Hawkins and Kenneth A. Coney, "Advertising and Differentiated Sex Roles in Contemporary American Society," Journal of the Academy of Marketing Science, 4 (No. 1 Winter 1976), 418-420.
- Kenneth Kemp, Least Squares Analysis of Variance: A Procedure, A Program, and Examples of Their Use. Kansas Agricultural Experiment Station, Contribution 168, Kansas State University, 1976.
- Jane Koester and Rose Adkins. Writer's Market '75, Cincinnati: Writer's Digest, 1974.
- George P. Morris and Edward W. Cundiff. "Acceptance by Males of Feminine Products," Journal of Marketing Research, 8 (August 1971), 372-4.
- Lee Sechrest, "Personality," In M. R. Rosenzweig and L. W. Porter (Eds.), Annual Review of Psychology, Palo Alto: Annual Reviews Inc., Vol. 27 (1976), 1-27.
- Janet T. Spence, Robert Helmuich, and Joy Stapp. "The Personality Attributes Questionnaire: A Measure of Sex Role Stereotypes and Masculinity-Femininity," Journal Supplement Abstract Service Catalog of Selected Documents in Psychology (Ms. No. 617), 4 (1974), 43.

John R. Stuteville. "Sexually Polarized Products and Advertising Strategy," Journal of Retailing, 47 (No. 2, Summer 1971), 3-13.

W. T. Tucker, "A Long Day of Discrepant Behavior," Proceedings, Fall Conference, American Marketing Association, 1976, 351-353.

Paul C. Vitz and Donald Johnston. "Masculinity of Smokers and the Masculinity of Cigarette Images," Journal of Applied Psychology, 49 (No. 3, 1965), 155-159.

Roland E. Wolseley, Understand Magazines, Ames, Iowa: Iowa State Univeristy Press, 1965.

ATTITUDE MATURATION: CHANGES IN RELATED
BELIEF STRUCTURES OVER TIME¹

Jerry C. Olson, Pennsylvania State University
Philip A. Dover, Dartmouth College²

Abstract

This paper presents some initial ideas about the dynamic changes which may occur within the belief structure associated with a stimulus concept. In particular, we focus on a phenomenon termed attitude maturation -- the observed tendency for the belief structure and overall attitude associated with a product to become increasingly related and consistent over time as knowledge about the product accumulates. An initial conceptualization of the attitude maturation phenomenon is offered along with several alternative explanations and these ideas are illustrated with data from a five-wave longitudinal experiment. The paper concludes with suggested theoretical perspectives derived from cognitive psychology.

Introduction

How are attitudes acquired? Are crystallized, stable attitudes formed "instantly" upon learning something of an object's existence and its characteristics? Or, do our attitudes develop (evolve or mature) more slowly as we accumulate information and cognitively consider that information? How long (how many experiences, how much information) does it take to form a stable, mature attitude -- and an associated stable, mature belief structure? Does the formation of beliefs through the acquisition of information directly and automatically lead to the formation of a stable (reliable) attitude that possesses clearly defined, stable relationships with associated salient beliefs and with other attitudes? Or, does this process take time and some degree of conscious, cognitive processing effort? This paper presents our initial thinking, accompanied by some illustrative data, regarding these and other questions involved with attitude formation, attitude change, and the dynamics of the relationship between attitude and belief structure.

The theoretical literature on attitude formation typically does not address such issues. Fishbein's attitude theory presents perhaps the clearest account of the attitude acquisition process (cf. Fishbein, 1963, 1967; Fishbein and Ajzen, 1975) and has been adopted or advocated by many consumer researchers (Ahtola, 1975; Cohen, Fishbein and Ahtola, 1972; Lutz, 1975; Olson and Mitchell, 1975). In this view, attitude, defined as unidimensional evaluation or affect, is a direct causal function of one's beliefs about associated concepts and one's evaluations of those perceived associations. Beliefs and their evaluative aspects are developed through learning principles of classical and instrumental conditioning and mediated generalization (Fishbein, 1967; Olson and Mitchell, 1975). Fishbein quantified the relationship between attitude (A) and belief structure (indicated by belief strength, b_i and evaluation, e_i) via the now familiar additive compensatory model:

¹This research was funded by a Research Initiation grant to the first author from the Pennsylvania State University. We are grateful to Rajesh Kanwar for his assistance with the data analysis and to James Bettman, Joel Cohen and Richard Lutz for helpful comments on a draft version of the paper.

²Associate and Assistant Professor of Marketing, respectively. This study and paper were completed while the second author was a doctoral candidate in Marketing at the Pennsylvania State University.

$$A = \sum_{i=1}^n b_i e_i \quad (1)$$

Implicit in most attitude research which takes this perspective is the assumption that a completely formed and internally consistent cognitive structure (here consisting of an attitude and its related belief structure) is formed essentially immediately upon processing whatever information is involved (cf. Lutz, 1975, 1977). However, the questions posed above challenge this assumption. To provide answers to these questions a dynamic, longitudinal research approach is required in which changes in attitude, belief structure, and their interrelationships are monitored over time for a given group of subjects. Unfortunately, the vast majority of consumer attitude studies are static and cross-sectional and thus do not address these issues. And, the relatively few longitudinal experiments that do exist (cf. Ginter, 1974; Winter, 1974) typically do not report the types of data analyses required for a study of cognitive dynamics and thus neglect the issues identified above.

Our purposes in writing this paper are several. First, we wish to present and briefly discuss some interesting conceptual issues regarding the dynamics of cognitive structure--from an attitude perspective. Where appropriate, specific issues will be illustrated with data from our own research. It should be noted that the data reported here are intended to be only heuristic and illustrative--not conclusive--as they are derived from a longitudinal experiment originally designed for another purpose. Second, we wish to identify critical theoretical issues that seem to warrant future research attention. Broadly speaking, our goals are to arouse interest in issues of attitude structure dynamics and to encourage research specifically directed at such issues.

Background

Definitions

The common tendency in attitude/cognitive structure research to use concepts and terms somewhat loosely suggests that specific, explicit definitions of the major concepts used in this paper would minimize semantic misunderstandings.

Attitude. Consumer researchers appear to be gradually reaching a consensus that the construct of attitude is best considered as the unidimensional evaluation or affect (favorable-unfavorable, good-bad, like-dislike) associated with an object, event, action, or concept. This simple definition was adopted for the present study as the most easily justified theoretically and the most useful definition from a measurement standpoint.

Cognitive Elements. The term cognitive elements as used in this paper refers to several major cognitive variables, including beliefs (b_i), the evaluative aspects of each belief (e_i), attitude toward the concept of interest (attitude toward the object, A_o , or act, A_{act}), and behavioral intention (BI, in consumer research usually operationalized as purchase intention).

Cognitive Structure. An implicit assumption in most cognitive research is that all of the cognitive elements associated with a stimulus object or event are organized together in a more or less consistent (in some sense) way and this organized set of cognitions may be

viewed as a structure. The most basic cognitive elements in a structure are the encoded representations of the stimulus object and its characteristics. These learned concepts are represented cognitively by symbolic codes assigned to the concept during the encoding process. Beliefs refer to the cognitive link between two concepts and belief strength is the perceived degree of probabilistic association between any two concept symbols (e.g., brand X and attribute A). Taken as a whole, the associations between these concepts may be considered to represent a belief structure as discussed below. The belief structure that is salient for a particular object/concept is thought to be related in a causal way to the overall attitude toward the concept of interest (see equation 1; Fishbein, 1967). In turn, attitudes (and perhaps other factors such as social norms) influence intentions to behave, which in turn are thought to exert a causal effect on overt behavior (see Ryan and Bonfield, 1975). Since concepts or codes, linking beliefs, evaluation/attitudes, and intentions are all cognitive elements, the overall pattern of their interrelationships may be considered as partially representing the cognitive structure for a particular concept such as a brand. However, because this paper is primarily concerned with attitudes, we shall deal mainly with belief structure interrelationships and belief-attitude relationships.

Belief Structure. The various salient and non-salient beliefs associated with an attitude concept can be considered as a structure of interrelated beliefs. That is, each belief concept is associated with the attitude concept and may also be "interconnected" with other belief concepts. The overall evaluation associated with a belief structure may be quantified by summing the cross products of belief strength and belief evaluation, $\sum b_i e_i$ (cf. Lutz, 1975; Olson and Dover, in press). This belief structure evaluation should influence (perhaps cause) the evaluation or attitude associated with the overall concept.

Method

The study which produced the data used here for illustrative purposes was designed to investigate another phenomenon--disconfirmation of expectations. The relevant methodology has been reported in sufficient detail elsewhere (Dover and Olson, 1977; Olson and Dover, 1976) and therefore is only briefly reviewed below.

Two treatment groups were randomly created from among 38 adult housewives. Group A (n = 20) received five exposures to information about a new, unfamiliar (actual) brand of ground coffee. The first three informational stimuli were three written, ad-like messages each of which stated (in different ways) that the coffee had absolutely no bitterness. No information about other coffee attributes were contained in any of the ads. Following the third message, consumers actually tasted a coffee sample intentionally prepared to be bitter and thus to disconfirm pretrial expectations. Then, to provide an extended period of product experience (and extensive product information) consumers received a free one-pound can of the coffee to try in their homes for about 10 days. The first four stages were each separated by approximately four days. Following each of the five stages, the major cognitive elements (associated with this brand of coffee) were measured. In sum, the five experimental stages for Group A can be seen as five opportunities for consumers to acquire information about a previously unknown product, with at least some of the information likely to be inconsistent with the remainder.

Consumers in Group B (n = 18) received no pretrial information regarding the coffee or its level of bitterness. These consumers merely tasted the purposefully bitter coffee and then took a one-pound sample home for

an extended trial period. Thus, the cognitive structures of group B subjects were measured twice. Consumers in this group had two opportunities to acquire information about the unfamiliar brand--both involving direct trial experiences with the brand.

For purposes of the present paper, this longitudinal experiment allows a close examination of (a) the dynamics of the belief structure-attitude relationship and (b) changes in the interrelationships within belief structure, as a function of multiple, sometimes inconsistent product information. The relationship between beliefs and attitude was modeled by the Ahtola (1975) adaptation of the Fishbeinian expectancy-value model (see Dover and Olson, 1977).³

Issues and Data

Magnitude of Experimental Effects

The experimental treatments, as expected, had significant effects on several cognitive elements, particularly on beliefs about the degree of coffee bitterness (see Olson & Dover, 1976, 1977). A substantial amount of "movement" over time was produced in the various cognitive variables by the exposures to the experimentally controlled information (see Figure 1, Dover and Olson, 1977). Our present interest, however, is with the dynamics (trends, shifts, or fluctuations) involved in the interrelationships between cognitive variables.

Major Interrelationships Among Cognitive Elements

One relationship of major interest to an attitude theorist involves the degree of correspondence between beliefs and overall attitude. As discussed previously, belief structure evaluation can be operationalized by the simple model, $\sum b_i e_i$. Column 1 in Table 1 contains the results pertaining to the validity of the basic Fishbein/Ahtola notion that belief structure is strongly and causally related to attitude ($\sum b_i e_i = A_0$).⁴ A steady increase⁵ in the strength of relationship between the belief structure index and attitude was obtained for Group A, increasing from an r^2 of .14 upon first learning something about the product to .61 after three weeks of varying experiences with the product. It is this effect that we have termed the attitude maturation phenomenon. The change in the attitude-intention relationship was even more dramatic, steadily increasing to a correlation of .95 ($r^2 = .90$) by the study's end. Group B

³This study adopted Ahtola's vector model, which separates belief strength from belief content, in order to more precisely monitor belief structure changes. In this model, a specific product attribute is reduced to a vector of discriminable levels or amounts of the attribute. Measures of belief strength and evaluation are taken for each attribute level. Summing the belief strength-evaluation cross products ($\sum b_i e_i$) within a vector yields an index of evaluation for that attribute. Summing across vector scores ($\sum \sum b_i e_i$) yields the composite evaluative index for the entire belief structure.

⁴Attitude toward the object (the coffee brand) was measured by the mean of three measures of evaluation -- good-bad, like-dislike, and high quality-low quality--each measured on 5-point bi-polar scales.

⁵We were unable to find a procedure for statistically testing differences between correlations of the same variables measured at different points in time for the same subjects.

TABLE 1
CORRELATIONS RELEVANT TO MAJOR THEORETICAL
RELATIONSHIPS BETWEEN COGNITIVE VARIABLES

Groups, and Experimental Stages	Theoretical Relationships	
	$\sum \sum b_i e_i = A_0$	$A_0 \approx BI$
<u>Group A (n = 20)</u>		
1. Ad A	.37	.54
2. Ad B	.39	.70
3. Ad C	.59	.76
4. Initial trial	.64	.93
5. Extended trial	.78	.95
<u>Group B (n = 18)</u>		
1. Initial trial	.48	.49
2. Extended Trial	.68	.63

Note: All correlations are significantly greater than zero ($p < .05$, one-tailed test).

also evidenced increases (although less dramatic) in the strengths of both relationships over the extended trial period (see Dover and Olson, 1977, for a more complete analysis and discussion of these results from an Expectancy value model perspective).

Interpretation. Despite the relatively small samples involved, at least compared to most correlational studies, increases of this magnitude and consistency in theoretically predicted relationships aroused our interest. A number of rather simple explanations for these dynamic effects were considered and tentatively judged unable to account for the effect (see Dover and Olson, 1977). For instance, increasing model correlations were not caused by a steady monotonic increase in the magnitude of the individual cognitive elements which might in turn produce increasing consistency between the cognitions. In fact, the variables fluctuated in magnitude across the experimental stages as would be expected given the conflicting information at stage 4 (see figures in Dover and Olson, 1977; Olson and Dover, 1976). However, the observed variations of the manipulated bit-terness belief, belief structure, attitudes, and intentions were relatively consistently related (see figure 1, Dover and Olson, 1977). That is, although cognitive changes were not monotonic over time, the change patterns for individual cognitive elements were similar.

An alternative explanation for the attitude maturation effect is that the observed increases in correlations were due to the repeated measurement of consumer subjects. However, this simple statement does not explicate precisely how the multiple measurements could have caused the correlation increases. For instance, it seems unlikely that subjects remembered their previous responses and/or consciously modified their responses to create a more consistent cognitive structure over time. The number and complexity of the measures and the relatively long intertrial time (at least four days) argue against this explanation. However, repeated, indeed any, measurements of cognitive variables certainly affect the measured cognitive structure and presumably also influence the relationships between cognitive elements. Thus, questions remain as to exactly what these measurement effects are and how they operate.

Belief Structure Dynamics

A more conceptually intriguing explanation for the observed attitude maturation effect concerns the dynamic changes that may have occurred within the belief structure underlying the product attitude. Using terms somewhat loosely at this point, the belief structure may become "clearer" (to the subject), more stable and reliable, more internally consistent--in other words, increasingly mature--and, as a consequence, more strongly associated with overall attitude. Such changes would not necessarily be evidenced in the absolute magnitudes or levels of specific cognitive variables, such as increasing belief strength, but rather might be more subtly represented by changes in the interrelationships between cognitive variables. In the following sections we discuss these possibilities and propose specific, theoretically based explanations regarding the belief structure changes, and illustrate their viability with data from the experiment described above.

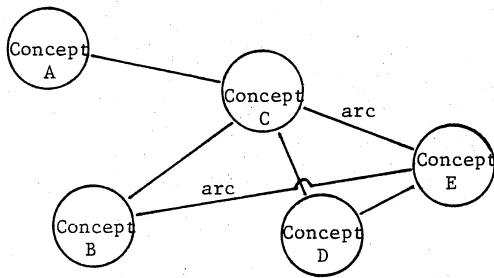
How is Information Structured and Stored?

It has been common in cognitive psychology and to a lesser extent in consumer research to discuss cognitive structure in terms of specific memory stores; e.g., sensory, short-term, intermediate, and long-term memories (see Chestnut and Jacoby, 1977 or Olson, 1977, for overviews). This approach led to rather mechanistic theories of information transfer between these structural "warehouses" via the "forklift trucks" of rehearsal or various retrieval processes. Currently these notions are being rejected for a less rigid concept of memory structure. In this view, memory or cognitive structure takes form as information is encoded and organized (cf. Estes, 1975; Hayes-Roth, 1977; Restle, 1974). That is, information enters memory and becomes a part of cognitive structure simply by having been encoded during initial processing activity.

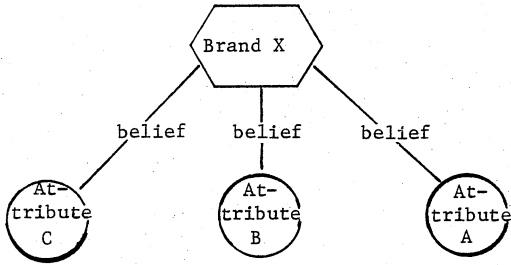
From this view, cognitive structure involves the content and organization of knowledge as stored in memory. Because of the focus on the meaning assigned to processed informational stimuli and the relations between that meaning and other, previously acquired knowledge, this aspect of cognitive structure or memory has been termed semantic memory (see Tulving, 1972). In sum, semantic memory concerns not the structural aspects of the storage system, but the structural characteristics of the knowledge that is represented in memory.

The traditional theoretical paradigm to explain knowledge representation in cognitive structure has been associationist theory (e.g., Anderson and Bower, 1973). Associative models of memory typically assume that knowledge is represented as networks of nodes and connecting arcs, where the nodes encode specific discriminable concepts and the arcs between nodes encode the perceived relationships between concepts (see Figure 1-a). Thus, cognitive structures can be modified by changing arcs (modifying the perceived association between two concepts) or by adding new concepts (and, of course, connecting arcs) to the cognitive structure.

Consumer attitude researchers will note the distinct similarity of these notions to their typical conceptualization of belief structures (Figure 1-b). Lutz (1975), for example, discussed strategies for modifying belief structures (and subsequently, attitude) in analogous terms. In their ubiquitous measurement of consumers' belief structures for products and brands, consumer researchers have in fact been investigating aspects of consumers' knowledge structure in semantic memory. However, the usual focus in consumer behavior/marketing has been on predicting attitudes from such belief structures rather than on describing the structural characteristics of the knowledge representations associated with a



a. Cognitive Structure in Associationist Terms



b. Cognitive Structure in Attitude-Belief Terms

FIGURE 1. ALTERNATIVE VIEWS OF COGNITIVE STRUCTURE

product concept.

Given this perspective, this paper can be seen as an initial, exploratory attempt to examine consumers' semantic structure per se, particularly focusing on the belief structure and the dynamics of its relationship with attitude. Following are several possible (and probably only partial) explanations for the attitude maturation effect. These ideas focus on changes within the structure of salient beliefs regarding the attitude object. Where possible, each notion is illustrated by data analyses from the study described above.

Belief Stability

One explanation for the observed maturation effects is that beliefs may not be well-formed or stable or "held with conviction" until several "experiences" with the belief object have occurred. There may be some problem in developing measures of this "stability" or "conviction" construct, since belief strength scores may not adequately assess such changes. That is, it seems possible that one could have a stable, clear belief that a particular concept is only weakly (or occasionally) related to another concept. In fact, belief strength scores for Group A did not increase (or decrease) consistently over the multiple stages (see Figures 1, 3, and 4 in Olson and Dover, 1976).

In the social and cognitive psychology literatures, researchers have taken independent measures of subjects' confidence in their belief ratings (cf. Wyer, 1973, 1974). Such confidence scores could be treated as direct evidence for the "clarity," stability, or degree of conviction of one's beliefs. In the present study, subjects were asked to rate their degree of confidence in the belief strength allocations for each of the five attribute

vectors.⁶ Table 2 presents the resulting five sets of confidence scores for each experimental stage. Confidence in beliefs about each attribute increased over the five stages for group A (all p 's < .10).⁷ Similarly, confidence increased over the two stages for group B for all belief vectors except bitterness (p 's < .10). However, group A's confidence after five exposures to product information was stronger than group B's confidence after two such exposures, for all attributes except the bitterness vector (p 's < .10). Thus, it would seem that a relatively steady increase in belief clarity, stability, or certainty occurred with successive exposures to information about the belief object. The confidence data, therefore, support the notion that an increasingly well-developed, stable belief structure develops over time as one becomes more knowledgeable or familiar with the relevant object.

Structural Relationships Between Beliefs

Another perspective to belief structure dynamics is provided by the notion that the interrelationships between the beliefs associated with an object/concept develop rather slowly and require time and experience before they become clearly defined and stable. This idea reflects the commonly-accepted tendency toward consistency among cognitive elements, but suggests that consistency may evolve over time with successive experiences which "activate" the belief structure. It seems logically reasonable that a well-formed, internally-consistent belief structure is more strongly related to overall attitude toward the belief object than a less well-developed structure.

Note, however, that the basis for consistency among beliefs has not been specified. Broadly speaking, two types of cognitive consistency have been investigated in past research -- consistency in terms of evaluative reactions and consistency based on perceived logical, meaningful, or semantic linkages between cognitive concepts (a review of this area is provided by Fishbein and Ajzen, 1975, pp. 143-156). For example, we may develop belief structures composed of beliefs that are increasingly similar in evaluative content--e.g., all or most of a consumer's beliefs about this coffee become more positive (or more negative). Alternatively, beliefs may be consistent in a meaningful, semantic sense; for example, one belief connotatively implies another belief. Stated differently, beliefs may be consistently related because they "fit together" on a logical basis--their meanings are consistent with one's previously learned expectations. For example, if one learns that a winter jacket is filled with goose down, one might also logically develop beliefs that the jacket is lightweight, expensive, and casual in style. In this case the consistency within the belief structure would be measured by the probabilistic relationships between belief concepts (belief strength) and not necessarily by increasing evaluative consistency. In summary, one might postulate a "need" or tendency to acquire internally-consistent cognitive structures, the evolution of which could occur over time as a function of processes "driven" by goals of evaluative or semantic consistency, or both.

⁶ Confidence in the 10-point belief strength allocation to the levels along each belief vector was measured on a 5-point scale with bi-polar labels not-at-all confident and very confident.

⁷ Due to the relatively small sample sizes (which decrease the power of statistical tests) and our greater tolerance for Type 1 errors and reluctance to commit Type 2 errors at this early state in our research, an alpha level of .10 was adopted for these analyses.

TABLE 2
MEAN CONFIDENCE IN BELIEF STRENGTH JUDGMENTS
FOR EACH ATTRIBUTE VECTOR

Groups and Experimental Stages	Product Attributes/Belief Vectors				
	Bitterness	Caffeine	Expense	Strength of Flavor	Consistency
Group A (n = 20):					
1. Ad A	3.70	3.35	3.65	3.70	3.45
2. Ad B	3.95	3.70	4.25	4.05	3.50
3. Ad C	4.15	3.75	4.10	4.05	3.80
4. Initial trial	4.40	3.40	3.95	4.35	4.05
5. Extended trial	4.55	3.95	4.10	4.60	4.20
Group B (n = 18):					
1. Initial trial	4.44	2.50	2.55	3.88	2.72
2. Extended trial	4.05	3.05	3.22	3.89	3.55

Evaluative Consistency. The following analyses concentrate on the relative changes in evaluative consistency among the belief elements in cognitive structure. Two indices of the evaluative consistency among the beliefs about the five product attributes were derived from the belief vector scores ($\sum b_{ie_1}$). These scores reflect the overall direction and degree of evaluation associated with each attribute. Estimates of internal consistency among these product attribute evaluation scores are provided by the coefficient alpha (α) index (see Nunnally, 1967, for a thorough discussion) and the average zero-order correlation between the five vector scores. Table 3 presents these two indices of evaluative consistency for both groups over the experimental stages. On both measures, group A evidenced a fairly dramatic increase in evaluative consistency after the second informational exposure, with relatively little increase in evaluative consistency thereafter. The evaluative consistency of belief structure for group B increased to a relatively low level over the two trial experiences.

TABLE 3
EVALUATIVE CONSISTENCY OF BELIEF STRUCTURE, BASED ON
CORRELATIONS BETWEEN BELIEF VECTOR SCORES ($\sum b_{ie_1}$)

Groups and Experimental Stages	Internal Consistency (coeffic. α)	Average Inter-item Correlation
Group A (n = 20):		
1. Ad A	.33	.09
2. Ad B	.57	.21
3. Ad C	.66	.28
4. Initial trial	.62	.25
5. Extended trial	.56	.20
Group B (n = 18):		
1. Initial trial	.03	.01
2. Extended trial	.29	.07

Yet another perspective on the evaluative consistency within a belief structure is provided by examining the relationship of each attribute vector ($\sum b_{ie_1}$) with the independent, direct measure of product attitude. Changes in these relations over time provide additional evidence for dynamic shifts in evaluative consistency within the belief structure. Table 4 presents the correlations between each belief vector evaluation score ($\sum b_{ie_1}$) and the direct rating of overall product attitude (A_0) for group A (recall that "no bitterness" was the product attribute belief most directly manipulated by the experimental treatments). By the study's end for group A, the evaluation associated with bitterness beliefs was the most strongly related to A_0 , $r = .78$ ($r^2 = .61$), followed by that associated with the amount-of-caffeine beliefs, $r = .50$ ($r^2 = .25$), and beliefs about the coffee's strength of flavor, $r = .47$ ($r^2 = .22$). The expensiveness and consistency vectors were only weakly related to overall attitude ($r^2 = .08$). A basically similar pattern of relationships was found for group B.

In sum, changes in evaluative consistency among salient beliefs seem to be somewhat related to the attitude maturation effect. It should be noted, however, that the present study probably attenuated the degree of evaluative consistency because the product information was not consistent over the course of the experiment. However, it does seem reasonable that as one's attitude structure "matures," one tends to acquire beliefs relatively consistent in terms of evaluation. A major problem with this explanation is in identifying sufficient and/or necessary amounts or degrees of evaluative consistency for a strong belief structure-attitude relationship. Threshold effects seem a distinct possibility. Such issues could be relatively easily explored in longitudinal correlational studies with larger samples.

Semantic Consistency. It is of interest to note that the three belief attributes most closely related to overall product attitude also seem closely related in a connotative or semantic sense as well. That is, an experienced coffee drinker might be expected to draw logical inferences about caffeine content and flavor strength after learning about a coffee's level of bitterness. There are probably relatively clear, semantically-based, expected linkages between these three attributes for experienced, knowledgeable consumers. In numerous studies, Wyer (cf. 1974, 1976) has quantified such semantic "connections" by asking subjects to rate the probability of one concept (or attribute level) given knowledge about the likelihood of another concept (or attribute level).

TABLE 4
CORRELATIONS OF BELIEF VECTOR SCORES (Σb_{ie_i}) WITH
DIRECT MEASURE OF OVERALL ATTITUDE (A_0)

Groups and Experimental Stages	Belief/Attribute Vectors				
	Bitterness	Caffeine	Expense	Strength of Flavor	Consis- tency
Group A (n = 20):					
1. Ad A	.26	.00	.45	.00	.26
2. Ad B	.15	.05	.19	.47	.36
3. Ad C	.28	.45	.47	.39	.39
4. Initial trial	.70	.30	.06	.60	.20
5. Extended trial	.69	.50	.29	.47	.29
Group B (n = 18):					
1. Initial trial	.43	.06	-.44	.48	.29
2. Extended trial	.29	.36	.17	.69	.15

Although such data was not collected in the present research, similar types of analyses could be done in consumer product judgments research, in effect yielding predictions of inferential belief formation, based upon conditional probability theory and analyses (see Olson, in press).

Summary. The confidence data (Table 2) and evaluative consistency data (Tables 3 and 4) support the inference that changes occurred within the product belief structure over the experimental stages. It seems logical that such changes in belief structure may cause (or at least partially influence) the increasing predictive validity for the Fishbein/Ahtola attitude model (see Table 1). Thus, the present analyses provide post-hoc support for the notion that the attitude maturation effect is due at least partially to changes occurring within the associated structure of salient beliefs. Clearly, however, before much confidence can be placed in these relationships and interpretations, similar measures and analyses must be included in subsequent research so that these initial findings can be replicated, or not.

Mediators of the Attitude Maturation Effect

To provide additional insight into the dynamics of the belief structure-attitude relationship as a function of multiple exposures to information, a final set of analyses was conducted to investigate possible mediators of such changes. This approach coincides with Chronbach's (1957, especially 1975) contentions that individual difference factors (variable characteristics of research subjects) tend to interact with experimental treatment variables in determining their effect on the dependent variable of interest, and thus should be explicitly examined. Similarly, Underwood (1975) has argued convincingly that researchers should make and test specific predictions of the mediating effects of individual difference variables in order to provide more rigorous tests of the theory of interest.

In anticipation of such use, several characteristics of the consumer subjects were measured in the study reported here (e.g., perceived product risk, coffee purchase volume). Because of their intuitive implications for belief structure dynamics, two of these variables were selected

TABLE 5
MEDIATING EFFECTS OF PRODUCT EGO-INVOLVEMENT AND SELF-CONFIDENCE ON THE ATTITUDE MATURATION
EFFECT AND ON THE INTERNAL EVALUATIVE CONSISTENCY OF BELIEF STRUCTURE

Group A Experimental Stages	Ego-involvement with Product						Product Specific Self-Confidence					
	$\Sigma b_{ie_i} = A_0$		Coeff. α		Ave. \bar{r}		$\Sigma b_{ie_i} = A_0$		Coeff. α		Ave. \bar{r}	
	Lo	Hi	Lo	Hi	Lo	Hi	Lo	Hi	Lo	Hi	Lo	Hi
	(n=10)	(n=10)	(n=10)	(n=10)	(n=10)	(n=10)	(n=10)	(n=10)	(n=10)	(n=10)	(n=10)	(n=10)
1. Ad A	.45	.27	.52	.28	.18	.07	-.11	.68*	.46	.28	.15	.07
2. Ad B	.53*	.21	.52	.65	.18	.27	.03	.49	.32	.71	.09	.33
3. Ad C	.79*	.22	.72	.39	.34	.11	.25	.73*	.46	.78	.15	.41
4. Initial trial	.84*	.56*	.69	.61	.31	.24	.67*	.64*	.75	.50	.37	.17
5. Extended trial	.79*	.81*	.58	.60	.21	.23	.30	.91*	.21	.69	.05	.31
Mean Correlation	.71	.46	.61	.52	.24	.18	.25	.73	.46	.62	.16	.26

* Correlation significantly greater than zero ($p < .05$, one-tailed test)

for presentation in this paper: (a) ego-involvement with the ground coffee product class, and (b) product specific self-confidence in evaluating and choosing alternative brands of ground coffee.

Ego-involvement. Consumers in group A were categorized either "higher" or "lower" in product ego-involvement based upon the relative number of coffee brands they sorted into the acceptable and neutral (noncommitment) regions for possible purchase.⁸ The 10 subjects who placed one-third or less of the familiar brands into the acceptance and noncommitment latitudes (or, alternatively, who rejected two-thirds or more of the brands) were deemed higher in product class involvement than the other 10 subjects.

Table 5 presents the relevant cognitive states and cognitive interrelationships for these "high" and "low" involvement subjects. Of greatest interest are the correlations between belief structure and attitude. Although the "high" and "low" involvement subjects yielded essentially the same belief structure-attitude relationship at the study's conclusion (r 's = .81 and .79, respectively), the more highly involved subjects reached that level of evaluative consistency more slowly. In fact, only after trial did the belief structures of the higher involvement group begin to become strongly related to overall attitude.

The explanation advocated for the lack of consistency between belief structure and attitude is that belief structures which are not well-formed, stable, or internally consistent may not be highly related to overall attitude. General support for this notion may be found by comparing the pattern of attitude model correlations for the "high" and "low" involvement subjects with the respective indices of evaluative consistency (see Table 5). Except for the data for the second stage (ad B), the evaluative consistency of belief structures is relatively low for those subjects with lower attitude-belief structure relationships, and vice versa for subjects with higher model correlations.

Product Specific Self-Confidence. A similar internal analysis was conducted by dividing group A subjects into two categories (n = 10 each) based on their relative self-confidence in making quality judgments of ground coffee.⁹ The product self-confidence variable also mediated dramatic differences in belief structure-attitude relationships (see Table 5). "Higher" confidence consumers had much greater consistency of belief structure and attitude than did subjects lower in confidence (r = .73 versus .25, respectively). However, the "internal evaluative consistency explanation" for this pattern of relationships was supported for only three of the five stages (2, 3 and 5).

Summary

The basic phenomenon of interest in this paper involves the cognitive structure changes termed attitude maturation--namely, the tendency for the salient belief structure associated with a stimulus concept to become in-

⁸This procedure is an adaptation of Sherif's ordered alternatives procedure for establishing involvement with an attitude issue. See Jacoby and Olson (1976) or Olson and Dover (1977) for detailed descriptions of the task and analysis.

⁹Self-confidence was assessed by a simple question which asked subjects to rate on a 5-point bipolar scale the degree to which they were "confident" in their ability to evaluate the quality of alternative brands of ground coffee.

creasingly predictive of the overall attitude towards that concept. Unfortunately, the present research design does not allow analysis of the unique effects of multiple measurement of cognitive structure (measurement and exposure to information were confounded), nor is it possible to completely eliminate the alternative explanation that the effect is caused by increasingly strong halo effects.¹⁰ Rather, this paper focuses on the heuristic and more interesting (to us) idea that cognitive structures, including the interrelationships among cognitive elements, develop and evolve toward evaluative and/or semantic consistency as one accumulates knowledge about the stimulus concept of interest. The illustrative data were generally consistent with this explanation, although there were exceptions. However, these analyses provide no explanations of the processes that cause such changes to occur.

The concluding section identifies three theoretical ideas derived from the cognitive psychology literature that may provide useful perspectives regarding the causal processes underlying cognitive structure dynamics. It is suggested that these theories can be combined to form an initial conceptual framework for explaining changes in cognitive structure, and can generate a number of interesting hypotheses to guide future research.

Evolution of Cognitive Structure

Very little has been written in the basic cognitive psychology literature regarding dynamic changes in cognitive structure. Recently, however, Hayes-Roth (1977) presented a clear theoretical description of the evolution or maturation of cognitive structure from initial formation to a well-formed state. Her theory is essentially consistent with our interpretations above. Briefly, Hayes-Roth stated that an organized cognitive structure begins with the establishment of encoded representations of elementary knowledge (cogits), which become stronger (more stable?) as experience accumulates. Basically this is the concept formation process (see Olson and Mitchell, 1975). As experience increases, these simple representations become cognitively linked to other representations and these associations become stronger. This is analogous to the belief formation process described by Fishbein (1967) and an attitude theorist might operationalize such changes as increases in either belief strength or confidence in belief ratings. Through repeated activations of these cognitions (caused by multiple exposures to the controlling stimulus), configurations of associated representations develop and become stronger. At some point a configuration may become so well-learned that it becomes "unitized," after which the entire cognitive structure may be treated as a single discrete memory representation or cogit (or perhaps a chunk, cf. Miller, 1956; Simon, 1974). Hayes-Roth emphasized that the development process for a cognitive structure proceeds "slowly," requiring repeated "activations" of the evolving structure.

Memory Schemata

The ideas regarding the evolution of highly organized cognitive structures are basically consistent with the heuristic concept of memory schemata (Olson, 1977). Norman and Bobrow (1975; Bobrow and Norman, 1975) based their conceptualization of memory schemata on the idea

¹⁰Note, however, that the two internal analyses yielded data that do not support the explanation of a steadily increasing halo effect. For example, the "high" involvement group produced strong belief structure-attitude relationships only after initial trial. And, the "high" confidence group produced relatively strong model correlations immediately upon the first informational exposure.

that the central goal of cognitive processing is the formation of a meaningful interpretation of the world; therefore, the sensory information available to a person must be encoded and organized relative to and in terms of some coherent framework of previously acquired knowledge (Norman and Bobrow, 1975, p. 119). Schemata are the large number of such structural frameworks of knowledge that a person has learned (developed or evolved) as his/her experiences have accumulated. A schema consists of the knowledge or encoded information about a concept as organized into a meaningful set of structural relationships. Thus, schemata not only contain representations of knowledge, but also "rules" of relationship between cognitive elements and even, for well-developed schemata, evaluation or decision rules for overtly responding to the schema concept. From a memory schemata perspective, conscious perception of a stimulus first involves the activation of the appropriate schema and then the interpretation or encoding of the stimulus "in light of" that schema. Thus, schemata may be considered to provide the framework of basic knowledge regarding a concept within which (and from which) initial encoding of incoming information (from ads or trial) relevant to that concept take place.

A number of interesting ideas and hypotheses regarding the attitude maturation effect can be derived from the schemata concept (see Olson, 1977). For instance, one might postulate that the development of a stable, internally consistent belief structure with a strong relationship to overall attitude is essentially equivalent to the formation of a well-developed schema. Following this reasoning, the belief structure encodes the interrelationships between knowledge concepts related to the attitude object. As the schema "matures," the "rule" for relating the knowledge structure to overall attitude toward the object concept develops and presumably gets stronger and more stable. One would expect a general expectancy-value model (if it is roughly similar to the true "rule") to show increasing correlations over time in such a situation. The present results are consistent with this line of reasoning. As one more example, one might expect consumers who are high in product self-confidence to possess more stable and well-developed schemata for that product. Such consumers, upon learning something about a new brand, should be able to quickly form a stable, well-developed belief structure for that brand, which in turn should be more strongly related to brand attitude. The internal analysis for self-confidence presented in Table 5 produced data consistent with this scenario. The belief structures for "higher" self-confidence consumers were much more strongly related to attitude than were those of the "lower" self-confidence consumers.

Depth or Level of Processing. A final notion from cognitive theory that appears relevant to the attitude maturation issue is "depth of processing." Craik and Lockhart (1974) introduced this concept to account for the common finding that memory performance (typically recall ability) is substantially enhanced when subjects perform cognitive "analyses" or tasks on the incoming stimuli that require their consideration of the meaning of the stimuli (cf. Jenkins, 1973). These results suggest that the encoding processes--that is, the specific cognitive operations carried out during initial processing of incoming information--are critical determinants of later memory performance. In particular, it seems that the degree of conscious attention focused on the meaning or semantic aspects of the stimulus is the critical type of encoding operation for improved memory performance.

Craik and Lockhart's basic notion, later modified somewhat (see Craik and Tulving, 1975), is that certain types or domains of encoding operations exist that can be characterized in terms of "depth," from less semantic to "deeper" levels of semantic processing. "Shallow" processing might concentrate on encoding physical characteris-

tics of the stimulus, while semantic processing would focus on various aspects of the meaning of the stimulus. The encoding operations applied to a stimulus do not necessarily proceed from sensory, shallow levels to deeper, more semantic domains. Rather, encoding processes may be elaborated at a particular level or domain, as indicated by the term "spread of processing."

Several implications for the attitude maturation phenomenon can be derived from the levels-of-processing concept. Perhaps most broadly, it would seem that the development of a well-developed, internally consistent belief structure (or schema) would require relatively "deep" conscious attention to the semantic meaning of incoming product information, over a series of exposures. That is, consumers would have to consciously consider the meaning of product information (e.g., the function of an attribute) and determine for themselves the value of that characteristic. Perhaps in the real world this degree of "semantic involvement" with product information, derived either from advertising or direct experience, is not common, especially at the time of exposure. But, in the present study, this level of conscious processing was essentially "forced" upon the subjects via the process of completing the extensive questionnaire regarding their beliefs, attitudes, and intentions about the product. That is, a subject had to engage in rather "deep" semantic encoding processes in order to answer the detailed belief questions regarding the various product attributes. By this reasoning, one would expect, *ceteris paribus*, that successive exposures to information and forced semantic processing should lead to the development of relatively consistent, stable, mature attitudes and associated belief structures.

Note that this predicted relationship between "depth of processing" and the formation of well-developed belief structures or schemata provides a theoretically-based and more satisfying account of the measurement effect explanation discussed previously. In one sense, the extensive postinformation measurements can be seen as manipulations of relatively deep levels of processing which, over time should create well-developed, stable schemata (belief structures). Future research could test this explanation (a) by including treatment conditions which receive the information of interest in the experiment, but receive no measurement of cognitive structure until the final stage, or (b) by forcing semantic analysis of the product information in ways other than responding to a questionnaire (e.g., tell someone else about the product).

Summary

The theoretical ideas presented here suggest that attitudes and associated belief structures evolve or mature over time as a function of repeated activations involving "deep" processing. The illustrative data cited were generally consistent with this hypothesis. The concept of attitude/cognitive structure maturation is theoretically interesting and, if verified, has clear practical implications for marketing research. For example, one should not be dismayed to find weak correlations for an expectancy-value attitude model applied to a new or unfamiliar product. Second, these ideas suggest that communication research might be focused on developing methods of enhancing cognitive structure "maturity." Finally, the maturation phenomenon is highly relevant to major marketing issues such as repetition effects of advertising (Mitchell and Olson, 1977) and the development of brand loyalty (Jacoby and Olson, 1976), among many others.

Many problems, of course, remain to be solved. Foremost among these are the clear conceptualization and development of valid measures of structure "stability/consistency/maturity." This paper presents several measure candidates, but future work must carefully examine the con-

vergent and divergent validities of these measures in a variety of situations. In contrast to the empirical problems, theoretical accounts of the maturation phenomenon are more easily provided. Several concepts from the cognitive theory seem useful in explaining these dynamic phenomena and in formulating future research questions. These include the Hayes-Roth (1977) theory of evolving cognitive structures, the Norman and Bobrow (1975) concept of memory as composed of numerous structural frameworks of organized knowledge called schemata, and Craik and Lockhart's (1972) notions of various types of encoding operations and their effects on information storage and memory.

As mentioned at the outset, our major goal was to stimulate interest in issues regarding changes in cognitive structure over time and as a function of specific informational manipulations. We hope this paper is an additional stimulus to the emerging interest among consumer researchers in cognitive structure dynamics.

References

- Ahtola, Olli T. "The Vector Model of Preferences: An Alternative to the Fishbein Model," Journal of Marketing Research, 12(February 1975), 52-9.
- Anderson, J. R. and Gordon H. Bower, Human Associative Memory (Washington, D.C.: Winston, 1973)
- Bobrow, Daniel G. and Donald A. Norman, "Some Principles of Memory Schemata," in D. G. Bobrow and A. M. Collins, eds., Representation and Understanding: Studies in Cognitive Science (New York: Academic Press, 1975)
- Chestnut, Robert W. and Jacob Jacoby, "Consumer Information Processing: Emerging Theory and Findings," in Arch G. Woodside, Jagdish N. Sheth and Peter D. Bennett, eds., Consumer and Industrial Buying Behavior (New York: Elsevier North-Holland, 1977)
- Chronbach, Lee J., "Beyond the Two Disciplines of Scientific Psychology," American Psychologist, 30(February 1975), 116-27.
- Chronbach, Lee J., "The Two Disciplines of Scientific Psychology," American Psychologist, 12(1957), 671-84.
- Cohen, Joel B., Martin Fishbein, and Olli T. Ahtola, "The Nature and Uses of Expectancy-Value Models in Consumer Attitude Research," Journal of Marketing Research, 9(November 1972), 456-60.
- Craik, Fergus I. M. and Robert S. Lockhart, "Levels of Processing: A Framework for Memory Research," Journal of Verbal Learning and Verbal Behavior, 11 (December 1972), 671-684.
- Craik, Fergus I. M. and Endel Tulving. "Depth of Processing and the Retention of Words in Episodic Memory," Journal of Experimental Psychology: General, 104 (September 1975), 268-94.
- Dover, Philip A. and Jerry C. Olson, "Dynamic Changes in an Expectancy-Value Attitude Model as a Function of Multiple Exposures to Product Information," in Barnett A. Greenberg and Danny N. Bellenger, eds., Contemporary Marketing Thought, 1977 Educators Proceedings (Chicago: American Marketing Association, 1977), 455-60.
- Estes, William K., "Structural Aspects of Associative Models for Memory," in Charles N. Cofer, ed., The Structure of Human Memory (San Francisco: W. H. Freeman, 1975).
- Fishbein, Martin, "An Investigation of the Relationships Between Belief About an Object and the Attitude Toward that Object," Human Relations, 16(1963), 233-40.
- Fishbein, Martin, "A Consideration of Beliefs and Their Role in Attitude Measurement," in Martin Fishbein, ed., Readings in Attitude Theory and Measurement (New York: Wiley, 1967)
- Fishbein, Martin and Icek Ajzen, Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research (Reading, Mass.: Addison-Wesley, 1975)
- Ginter, James L, "An Experimental Investigation of Attitude Change and Choice of a New Brand," Journal of Marketing Research, 11(February, 1974), 30-40.
- Hayes-Roth, Barbara, "Evolution of Structures and Processes," Psychological Review. 94(May 1977), 260-78.
- Jacoby, Jacob and Jerry C. Olson. "An Attitudinal Model of Brand Loyalty: Conceptual Underpinnings and Instrumentation Research." Paper No. 159, Purdue Papers in Consumer Psychology, Department of Psychological Sciences, Purdue University, 1976.
- Jenkins, James J. "Remember that Old Theory of Memory? Well, Forget It!" American Psychologist (November 1974), 785-95.
- Lutz, Richard J, "Changing Brand Attitudes Through Modification of Cognitive Structure," Journal of Consumer Research, 1(March 1975), 49-59.
- Lutz, Richard J, "An Experimental Investigation of Causal Relations Among Cognitions, Affect and Behavioral Intention," Journal of Consumer Research, 3(March 1977), 197-208.
- Miller, George A, "The Magic Number Seven, Plus or Minus Two: Some Limits in Our Capacity for Processing Information," Psychological Review, 63(March 1956), 81-97.
- Mitchell, Andrew A. and Jerry C. Olson, "Cognitive Effects of Advertising Repetition," in William D. Perrault, Jr., ed., Advances in Consumer Research: Vol. IV (Atlanta: Association for Consumer Research, 1977), 213-20.
- Norman, Donald A. and Daniel G. Bobrow, "On the Role of Active Memory Processes in Perception and Cognition," in Charles N. Cofer, ed., The Structure of Human Memory (San Francisco: W. H. Freeman, 1975)
- Nunnally, Jum, Psychometric Theory (New York: McGraw-Hill, 1967).
- Olson, Jerry C., "Theories of Information Encoding and Storage: Implications for Consumer Research." Paper presented at the Conference, "Effects of Information on Consumer and Market Behavior," Carnegie-Mellon University, May 20 and 21, 1977.
- Olson, Jerry C., "Inferential Belief Formation in the Cue Utilization Process," in H. Keith Hunt, ed., Advances in Consumer Research: Vol. 5. Association for Consumer Research, in press.
- Olson, Jerry C. and Philip A. Dover, "The Effects of Expectation Creation and Disconfirmation on Belief Elements of Cognitive Structure," in B. B. Anderson, ed., Advances in Consumer Research: Vol. III (Cincinnati, Ohio: The Association for Consumer Research, 1976), 168-175.
- Olson, Jerry C. and Philip A. Dover, "Cognitive Effects of Deceptive Advertising," Journal of Marketing Research, in press.
- Olson, Jerry C. and Andrew A. Mitchell, "The Process of Attitude Acquisition: The Value of a Developmental

Approach to Consumer Attitude Research," in Mary J. Schlinger, ed., Advances in Consumer Research: Vol. II (Chicago: Association for Consumer Research, 1975), 249-264.

Restle, Frank, "Critique of Pure Memory," in Robert L. Solso, ed., Theories in Cognitive Psychology: The Loyola Symposium (Potomac, Md.: Lawrence Erlbaum, 1974)

Ryan, Michael J. and E. H. Bonfield, "The Fishbein Extended Model and Consumer Behavior," Journal of Consumer Research, 2(September 1976), 118-37.

Simon, Herbert A., "How Big is a Chunk?" Science, 183 (February, 1974), 482-8.

Tulving, Endel, "Episodic and Semantic Memory," in Endel Tulving and W. Donaldson, eds., Organization of Memory (New York: Academic Press, 1972)

Underwood, Benton J, "Individual Differences as a Crucible in Theory Construction," American Psychologist, 30 (February 1976), 128-34.

Winter, Frederick W, "The Effect of Purchase Characteristics on Postdecision Product Reevaluation," Journal of Marketing Research, 11(May 1974), 164-171.

Wyer, Robert S, "Category Ratings as 'Subjective Expected Values'; Implications for Attitude Formation and Change," Psychological Review, 80(1973), 446-67.

Wyer, Robert S., Jr., Cognitive Organization and Change: An Information-Processing Approach (Potomac, Md.: Lawrence Erlbaum, 1974.

Wyer, Robert S., Jr., "Effects of Previously Formed Beliefs on Syllogistic Inference Processes," Journal of Personality and Social Psychology, 33 (March 1976), 307-16.

CARRYOVER EFFECTS OF CORRECTIVE ADVERTISING

Alan G. Sawyer, The Ohio State University
Richard J. Semenik, University of Utah

Abstract

A laboratory experiment tested based on communication theory and empirical research, about the delayed or carryover effects of corrective advertising. The results of the tested corrective advertisements present evidence that one exposure of a corrective advertisement may not influence brand beliefs as would be hoped by public policy makers. Problems with the employed longitudinal design are discussed and suggestions for future research are offered.

Introduction

The effects of corrective advertising have been examined in several published laboratory experiments by consumer researchers. The general conclusion that may be gained from this literature is that one exposure of a corrective advertisement can, under certain circumstances, have the desired negative effect on target beliefs and attitudes. This paper examines the persistence of effects of corrective advertising over time. The thesis of this paper is that any obtained immediate effects of a corrective ad are likely to rapidly dissipate over time. The experiment described in this paper tests this proposition by including measures delayed up to eight weeks after the exposure of a corrective ad. In addition, it includes several factors which provide a communication situation for the tested corrective advertising more externally valid than some past research in this area.

After first discussing the theoretical and empirical rationale for the predictions about the carryover effects of corrective advertising, this paper will review four past experiments about corrective advertising. Following that review, the methodology of this experiment will be described and contrasted with the methodology of the previous research in this area. Although the described results fail to support the predictions about the immediate and delayed effects of corrective advertising, it is hoped that this paper will encourage other researchers to examine delayed as well as immediate effects of advertisements and other marketing communications.

Hypothesized Decay of Corrective Advertising Effects¹

Figure 1 shows the hypothesized results of an experiment which measures beliefs toward an established product both immediately before and after exposure to a corrective advertisement for the product and also several weeks after exposure. Although the effects of the corrective ad (point 2) might immediately neutralize or even reverse the previous cumulative effects of the deceptive advertising campaign (point 1), the belief based on deceptive prior information would reassert itself with time (point 3). Although no previous corrective advertising research has examined carryover effects, several other areas of research about the effects of attitudes over time lend strength to the above prediction. Especially relevant are two areas of investigation of carryover effects: 1) the general nature of attitude decay and 2) the relative

¹Space prevents a full discussion of the literature discussed in this section. For more details, see Sawyer (1976) and Sawyer and Ward (1976).

persistence effects of discounted or qualified messages.

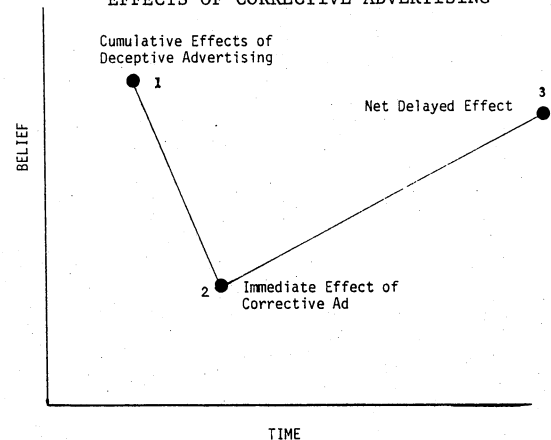
Attitude Decay

If, as some researchers believe (e.g., Miller and Campbell, 1959), decay of attitude change resembles the negatively accelerated forgetting curve, a prediction of Jost's Second Law of Learning (Jost, 1897) is very relevant to the corrective advertising situation. Figure 2A illustrates the prediction that, for two stimuli equal in strength but of different age at a given point in time (t_1), the effect of the older of the two will decay less rapidly - reaching some asymptotic level higher than that of the newer campaign. Figure 2B shows the same curves as in Figure 2A except that curve D shows the hypothesized "positive decay" of any negative effects of corrective advertising. A major assumption in Figure 2B is that, whereas a brand belief rated highly probable will decay over time "down" toward some neutral point or mean probability for all competing brands; a belief rated highly improbable will decay "up" towards the same point of neutrality or mean probability for all brands. A copy test may misleadingly indicate at time t_1 that a corrective ad (C) is effective at reversing the effects of a deceptive campaign (D). Over time, however, the effects of the much newer and less often exposed corrective ad is likely to decay more rapidly and result in a belief much more positive than the immediate effect.

Discounted Messages

The "discounting cue hypothesis" theorizes that, if there is some aspect of a message that causes the message to be initially discounted or resisted, the passage of time may lead to a dissociation of the inhibiting cue from the original message so that the potential attitude change impact of the persuasive content may gradually take effect (see Gruder, et al., 1974). The combined impact of a positive (deceptive) message followed by a negative (corrective) message can be viewed as a discounted cue situation where the corrective ad discounts the previous positive advertising claim(s). The results of each of the three discounting cue experiments, which examined the delayed effects of messages highly analogous to corrective advertisements, are consistent with the prediction graphed in Figure 1.

FIGURE 1
HYPOTHESIZED IMMEDIATE AND DELAYED
EFFECTS OF CORRECTIVE ADVERTISING



Weiss (1953) presented four exposures of a list of eight statements about the effects of smoking on health. Each statement was followed by the words "true" or "false" (e.g., "Studies have shown that smoking is an important factor in producing cancer of the lungs and throat. True"). Immediately after these four learning trials which were designed to produce anti-smoking beliefs, subjects in the discounted message group were also presented with a brief countercommunication in which they were told that more recent smoking experiments conclusively contradicted the findings of the research studies on which the statements were based and that they should be very skeptical about the truth of the original statements. Belief was then measured either immediately, three weeks, or six weeks after the messages.

FIGURE 2A
RELATIVE EFFECTS OF OLDER, MORE REPEATED
ADVERTISEMENT (A) COMPARED TO NEWER COMPETITOR AD
(B) WITH NEGATIVELY ACCELERATED ATTITUDE DECAY

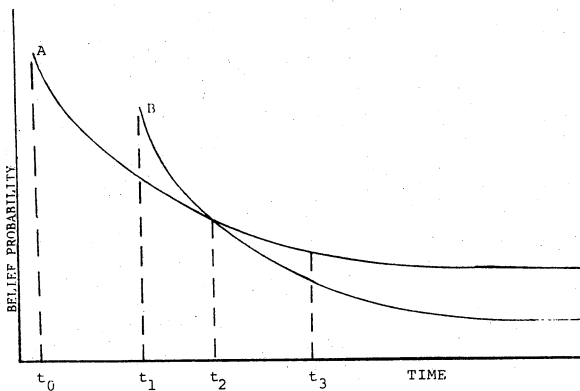


FIGURE 2B
RELATIVE EFFECT OF DECEPTIVE ADVERTISING CAMPAIGN
(D) AND NEW CORRECTIVE AD (C) WITH NEGATIVELY
ACCELERATED ATTITUDE DECAY

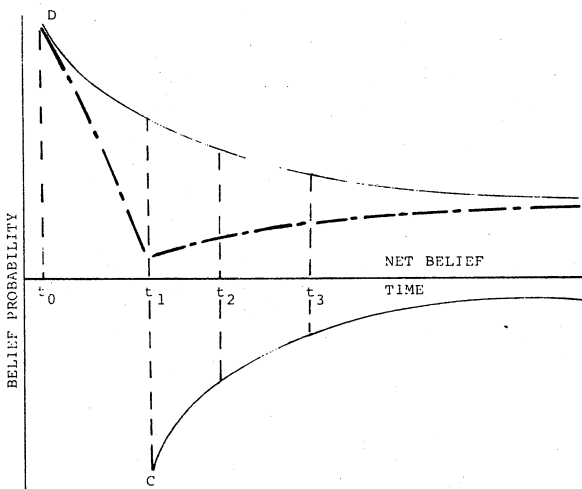
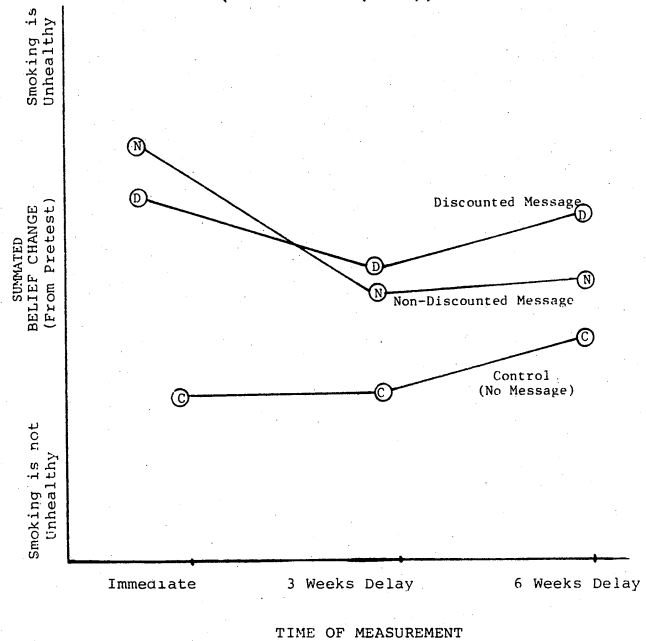


Figure 3 shows the effects of the discounted message condition which is directly analogous to a situation of a repeatedly exposed, deceptive advertising campaign advocating that smoking is unhealthy followed by a corrective advertisement to the contrary viewpoint. As would be desired by public policy makers, the discounted message produced initially lower beliefs about the unhealthy effects of smoking than the non-discounted message. After six weeks, however, the relative effectiveness of the discounted (corrective) message condition was

reversed from the immediate situation and produced significantly higher beliefs that smoking was unhealthy than the non-discounted condition. Ironically, the condition directly analogous to corrective advertising was even less effective in maintaining "correct" beliefs than the "uncorrected" messages when the effects were measured over time.

FIGURE 3
THE EFFECTS OF DISCOUNTED AND NON
DISCOUNTED MESSAGES ON BELIEFS
(from Weiss (1953))



A recent experiment by Gruder, *et al.* (1974) found not only significantly less negative attitude decay in the discounted condition compared to a non-discounted message but also found a significant positive slope of attitude over time for the discounted message. Subjects were exposed to two exposures of either one of two messages - one arguing against a new law allowing right turns on red lights and the other arguing that a four-day work week decreases worker satisfaction. The discounted message condition also included a following message that cautioned the subject that, since the selection of the previous message for the experiment, it had been learned that the message's arguments were not true and had positively been refuted as "inaccurate and, frankly, wrong." Attitude change was measured both immediately and five weeks after the messages.

Not only were the slopes different as predicted, but the attitudes produced by the discounted (corrective) messages actually became more positive over the five week period. Although this increased attitude did not exceed the attitude produced by the message only condition, the large differences in slopes add further support to the prediction that a corrective ad is likely to be ineffective in producing enduring negative effects to eradicate the residual attitude based on deceptive information.

In the corrective advertising used in the Profile and Ocean Spray cases, the employed format was essentially that of a qualified argument where the messages were positive in overall nature but contained corrections of misimpressions that might have been formed by past advertisements. Bartlett (1932) predicted that, for a message arguing for one side of an issue but also presenting some qualifications and reservations against the main arguments, the qualifying details would be forgotten more quickly than the main conclusion. Because of this

difference in forgetting, the qualified message would show a delayed action effect on beliefs with more change after a passage of time than immediately following the message.

Papageorgis (1963) tested Bartlett's hypothesis about the relative persistence of belief change of qualified and unqualified messages. Two messages low in ego involvement--one advocating that life expectancy was longer in cities than in rural areas, and the other arguing that there had been few technological advances in oil deposit location practices--were presented. A subject received the unqualified message for one topic and the qualified version for the other. The qualified message conditions contained the same arguments but stated much more equivocally with many reservations about the validity and generality of the contained arguments. Content recall and beliefs were measured immediately, two days, fourteen days, and forty-one days after the message.

The recall results were nearly exactly as predicted. For the qualified messages, recall of the main conclusions barely declined over forty-one days whereas there was substantial decay in the recall of the qualifications. As predicted, the decay in belief produced by the qualified messages was less than that of the unqualified messages. Although this difference averaged across both message topics was not statistically significant, the predicted relative advantage for delayed belief for the qualified message was statistically significant for the life expectancy issue. In other words, the qualified messages, which were quite similar in format to corrective advertisements, produced beliefs that, although initially lower as would be hoped by public policy makers, tended over time to revert toward the position held by subjects who viewed only unqualified messages.

Immediate Effects of Corrective Advertising

Four laboratory experiments on the attitudinal effects of corrective advertising have all indicated that the effects of deceptive ads tended to be at least partially neutralized by one corrective ad. Hunt (1973) found that a print form of a corrective ad about allegedly false claims by Standard Oil of California produced a more negative attitude about Chevron than a group which viewed only the deceptive ad. Similar negative effects of a corrective ad resulted whether the copy was a detailed, research-based refutation of the truthfulness of the original claims or a general statement that the original ad was false and deceptive, and whether the source of the ad was Standard Oil, a government agency, or an independent consumer organization. The explicit, detailed corrective ad appeared to be especially effective.

Kassargian, Carlson, and Rosin (1975) examined the effects on product image of a corrective print ad for a motorcycle safety helmet. This corrective advertisement, which was sponsored by a retailer, was effective in reversing the positive effects of the original deceptive ad on several general dimensions of product image. However, the negative effects of corrective advertising did not generalize to the retailer.

Dyer and Kuehl (1974) reported slightly more equivocal results of fictitious corrective advertisements for nationally advertised products. One corrective ad was a print ad about a diet soft drink and the other was a radio ad about a suntan lotion. Only when the F.T.C. was indicated as the source of the corrective message was the corrective ad for the suntan lotion effective in decreasing intentions to buy the product. When the source was the manufacturer of the product, the suntan lotion corrective ad was ineffective. Neither source was effective in decreasing purchase intention in the

radio corrective ad for a diet soft drink.

Mazis and Adkinson (1976) tested four advertising conditions for a well-known mouthwash brand currently negotiating with the F.T.C. about a corrective advertising order. Two corrective ads stressed that the brand was ineffective at preventing colds or sore throats but helped to kill germs. In addition, one of the corrective ads mentioned that the admission about the cold prevention ineffectiveness had been required by the F.T.C. One positive, non-corrective ad stressed the germ-killing effectiveness, and the other positive-only ad emphasized only the breath freshening qualities without making any germicidal claims. The results showed that aided recall of the corrective ads was higher than of the non-corrective ads. Analysis of variance, controlled for brand usage, found that the target belief about the probability that the brand was effective at preventing colds or sore throats was significantly lower after exposure to the corrective ads. A similar decrease was found for the non-target belief about germ-killing effectiveness. Not only was the target belief strength lowered, but the evaluation of the goodness of that product attribute was also lowered by corrective advertising. No effect of corrective advertising on overall attitude toward the brand was found.

Methodology

Table 1 classifies the four past research studies according to several important methodological dimensions. These areas include the research method, advertised brand, corrective advertising conditions, and measurement details. In addition to including the time factor, this experiment differed from the previous ones in several important respects.

Advertised Brand

The product which was the focus of the corrective advertising in this study was Listerine mouthwash, which is currently under negotiation for corrective advertising with the F.T.C. This product met several important criteria presented by Sawyer (1977). First, the product was a real brand familiar to the subject population. Although research using fictitious or unknown brands as stimuli is "clean", it is also irrelevant to the problem of whether and how corrective advertising can accomplish its goal of erasing the residual of a deceptive campaign. It is highly unlikely to have a strong residual belief about a fictitious or unfamiliar brand. Thus, it is not surprising that, in the studies which showed merely one exposure of a "deceptive" ad to subjects who had never seen that ad before, one exposure of the "corrective" ad was able to negate its effects (Hunt, 1973; Kassargian *et al.*, 1975).² Such a design can be viewed as a primacy-recency experiment which is not appropriate to the research question of whether corrective advertising can overcome and reverse the residual of a familiar, highly repeated deceptive campaign. In the study reported here, the use of a well-established ad campaign as the "deceptive" advertising permitted a natural experimental design (Campbell, 1969) in which the deception was already established or, at least, easily reestablished with a reminder exposure of the deceptive advertisement. Also, the fact that subjects may have been previously exposed to the publicity of the court case involving the mouthwash product and the F.T.C. added to the external validity of the research.

²The Chevron F-310 ad in Hunt (1973) was not familiar to the Iowa college student subjects in Hunt (1973). The motorcycle helmet ad which was sponsored by a New York City retailer had not been exposed to the Maryland residents who were subjects in the Kassargian, *et al.* (1975) study.

TABLE 1

Summary of Past Research on Corrective Advertising

	Research Method	Advertised Brand	Brand advertised and sold in local area?	Corrective Advertising Copy Conditions	Advertising Medium	Subjects	Beliefs on Deceptive Dimension Measured?	Measurement over time?	Other Independent Variables	Major Results
Hunt (1973)	Lab	Chevron F-310 gasoline	No	1) explicit--point-by-point corrections of deceptive claims 2) general--statement that past ads were "misleading, deceptive, and untrue" 3) control	Print	College students	No (affect toward brand)	No	Type of inoculation by advertiser in deceptive ad (supportive, refutational, none)	Corrective ads led to significant negative effect; explicit ad most effective--especially with refutational inoculation defense
Dyer and Kuehl (1974)	Lab	Diet soft drink; Suntan lotion	Yes	1) high strength explicit corrective language and disclosure of past deceptive advertising 2) low strength explicit corrective language but vague reference to "past misleading impressions" 3) control	Radio; Print	College students	No (aided brand recall; other brand beliefs; purchase intention; perceptions of both advertisement and company image)	No	Source of corrective ad (F.T.C.; offending company)	Purchase intention decreased by corrective ad only in high strength message; F.T.C. source condition led to less trustworthy corporate image
Kassarjian, Carlson, and Rosin (1975)	Lab	Motorcycle safety helmet	No	1) weak corrective ad sponsored by retailer admitting lack of accredited research substantiation 2) control	Newspaper	College students	No (ad recall; 15 general beliefs about both brand and retailer)	No	Ad exposure; natural (ad unobtrusively placed in five page newspaper) and forced (focused attention)	Corrective ad led to more negative image profile of brand but not of retailer; profiles did not differ between ad exposure conditions
Mazis and Adkinson (1976)	Lab	Mouthwash	Yes	1) corrective copy about cold prevention capability plus positive germ-killing appeal (F.T.C. source) 2) same corrective copy as condition 1 (company source) 3) non-corrective copy--only positive germ-killing appeal 4) non-corrective copy--only positive, non-germ killing appeal	Radio	College students	Yes (also evaluation of goodness of rated product dimensions; overall brand attitudes; brand usage; ad recall)	No	Control for whether subjects used test brand	Corrective ads decreased target belief about "cold prevention" capability and also belief about germ-killing capacity; evaluation of cold prevention also lower for corrective ads; no effect of ad source or brand usage on beliefs
This study	Lab	Listerine Mouthwash	Yes	1) non-corrective, positive only 2) negative only 3) two-sided, positive first 4) two-sided, negative first	T.V.	Adults	Yes (also importance of each of eight belief dimensions)	Yes	whether one exposure of deceptive ad was exposed prior to corrective ad	

Second, the brand advertised an explicit claim which was believed by the general population. Such believability was ascertained by a pretest--using belief measures described below. Listerine was rated as significantly ($p < .05$) more probable than Scope, Lavioris, or Cepacol to possess the attributes of "prevents colds", "relieves nasal congestion", and "reduces colds severity"; more likely than Scope or Lavioris to "kill germs"; and less probable than the other competitors to "have appealing taste."

Finally, the advertising claim refuted in the corrective advertising was not telecast after the initial laboratory exposure to the corrective advertising. After early May of 1976 when this experiment was begun, Listerine did not use its campaign which stressed the (allegedly false and deceptive) claim of preventing colds and reducing the severity of their side effects. The former campaign was replaced by a new campaign which stressed an appeal based on germ killing and breath freshening effectiveness.

Ad Copy and Medium

Five 30-second TV commercials comprised the treatment stimuli. These ads, including one deceptive message, three versions of a corrective message, and one non-deceptive positive message, featured a professional actress and were filmed in color by the Photography and Cinema Department of Ohio State University. The advertisements were buried within a half-hour video tape of a television program. The copy for the advertisements employed the recommended corrective copy offered by the Federal Trade Commission for the brand (FTC vs. Warner-Lambert, 1974, p. 102), as well as copy devised by the researchers. It should be noted that the advertisements were of such quality that none of the subjects doubted the authenticity of the advertisements.

Subjects

Unlike past studies, subjects were not confined to a

college student population. Of the 142 total subjects (122 experimental and 20 control), 50 were obtained from random mail solicitation and 113 subjects from a commercial marketing research firm. To give a general profile of the laboratory subjects, the average age of the participants was 43.1 years and they had an average household income of between \$12,500 and \$15,000. The sample included 29% males and 71% females; 32.7% of the subjects had college degrees; and only 15.6% were currently students.

Communication Measures

There have been two major problems with the cited past research on corrective advertising. It has, in our opinion, sometimes employed improper measures of effectiveness, and, as mentioned previously, it has measured only the immediate effectiveness.

If the FTC's goal for corrective advertising is to eliminate the residual of deception, then that residual must be defined. Advertising can attempt to achieve several communication goals including awareness, beliefs, attitudes, and intended or actual behavior. Concerning attitudes, an advertiser can try to influence either consumers' beliefs about some attribute(s) of its brand or competing brands or the evaluation consumers place on various attributes of brands in the product class in question. (Boyd, et al., 1972). It is very important to distinguish among a belief about how a particular brand rates on a particular attribute or dimension, an evaluation of that attribute, and a general affective attitude toward that brand (see Fishbein and Ajzen, 1975).

If an advertiser makes a deceptive claim about a particular attribute, the most direct effect would be to change consumers' beliefs about that attribute (e.g., Lutz, 1977). Of course, the ad may also have secondary effects on other beliefs or, through the changed beliefs, affect attitudes, purchase intention, or behavior. However, in the opinion of the authors, the prime criterion of impact of the deceptive claim is the change in the belief about that dimension and not anything else (see also Olson and Dover, 1978, Wilkie and Gardner, 1974). If belief is the prime residual of deceptive advertising, corrective advertising must focus on the particular dimension in question and change that belief to coincide with the ascertained facts.

In this study, informal interviews and pretests prompted the use of eight attribute dimensions of the mouthwash product category. Thus, the questionnaire measured eight beliefs about the advertised brand as well as three other competing brands and the perceived importance of those attributes. Most critical to the goals of this experiment, these measures were taken not only immediately following exposure of the corrective advertising but also two, four and eight weeks later.

Experimental Procedure and Design

Upon arrival at the laboratory, all participants were informed they would be participating in a study regarding advertising and general social issues and were asked to complete a questionnaire concerning attitudes toward advertising in general, attitudes toward government regulatory activities, product usage, and demographics.

The practice of some past experiments (e.g., Hunt, 1973, Kassarjian, et al., 1975) which included one exposure of the deceptive ad shortly prior to the exposure of the corrective ad seems low on external validity when, in an actual campaign, these ads would be widely separated. On the other hand, inclusion of an exposure of the deceptive ad might be necessary to trigger the deceptively-based belief. In order to test the impact of this factor, exposure to a deceptive ad closely preceding exposure to a corrective ad was experimentally manipulated. Half the

experimental participants were exposed to a deceptive advertisement, which argued that the product was effective in reducing colds and sore throats; the other half viewed an ad for an entirely different product. Following this exposure, half of each of these experimental groups was given a questionnaire assessing attitudes toward the firm, attitudes toward several mouthwash brands, and product performance beliefs. The other half was given a "bogus" questionnaire measuring attitudes toward social issues. This procedure was followed to determine whether any learning effects were present for subjects using the same rating scales both before and after the corrective messages. (Subsequent analysis (Semenik, 1976) revealed no effects of this pretest, so all subjects were combined into one analysis).

During the next phase of the procedure, all subjects were informed that they were to view a program aired on network television concerning the solid waste disposal problem. On a video tape monitor and playback unit, subjects watched a seventeen minute program on this topic with the appropriate treatment stimuli contained as part of the program.

Depending upon the treatment group to which they had been randomly assigned, experimental participants viewed one of the four corrective advertisements contained in the program (see Semenik (1976) for the exact copy of each): Corrective Version 1: One-sided positive appeal only. Since only favorable product attributes were communicated (deleting the deception in the previous advertising for the firm), this was not a corrective ad. This ad emphasized the breath freshening qualities. This version represents a possible outcome of a "cease-and-desist" order issued by the FTC. Corrective Version 2: One-sided negative only. This ad explicitly stated that the product was not effective at reducing colds and sore throats. A firm might be required to run such an advertisement in which only the previous deception is explained and the firm is not able to present favorable product attributes. Corrective Version 3: Two-sided, positive first. This communication related positive product attributes first and then revealed the deception of previous advertising. Corrective Version 4: Two-sided, negative first. The firm first presented unfavorable information (disclosure of deception) and then presented positive product attributes.

Following exposure to one of the four corrective advertisements, subjects rated, as a diversion from the true purpose of the experiment, several open ended questions about the effectiveness of the television program. Subjects were then asked to complete another questionnaire regarding attitudes toward the brand, attitudes toward the firm, and product performance beliefs. Upon completion, they were each paid \$2.00 and were told that they might be contacted again as part of their paid participation.

The experimental design was a 2 x 3 x 2 x 4 mixed, factorial design with order of measurement (immediately after ad exposure or delayed) as a within-subjects factor and length of delay (two, four, or eight weeks), prior exposure to original deceptive commercial or not, and corrective ad copy (positive-only; negative-only; two-sided, positive side first; and two-sided, negative side first) as between-subjects factors. Whereas all subjects were measured immediately, subjects in each ad copy-exposure condition were randomly assigned to one of the three delay conditions. Either two, four, or eight weeks later, they received a mail questionnaire which, after reminding them of their earlier participation, asked them to answer a "few last questions." These questions included belief and importance ratings identical to the previous immediate-past measures. Return rates were remarkably (and quite uniformly) high over the three delay conditions with 90.9%, 89.5% and 81.3%, respectively.

In addition to the experimental groups, another group (n=20) acted as an immediate post-measure control group. This latter group saw the same television show and filled out the same questionnaire but did not see any experimental ads. Regretably, a clerical error caused an omission of delayed questionnaires sent to the control group. Table 2 presents the experimental design and sample sizes.

TABLE 2
Experimental Design and Cell Sample Size

Also Exposed to Deceptive Ad				
Corrective Ad Copy Condition	Positive Only	Negative Only	Two-sided Positive First	Two-sided Negative First
Order of Measurement	Imm/Delay	Imm/Delay	Imm/Delay	Imm/Delay
Delay Condition:				
Two Weeks	5	5	6	5
Four Weeks	4	4	6	4
Eight Weeks	6	3	5	4
Not Also Exposed to Deceptive Ad				
Corrective Ad Copy Condition	Positive Only	Negative Only	Two-sided Positive First	Two-sided Negative First
Order of Measurement	Imm/Delay	Imm/Delay	Imm/Delay	Imm/Delay
Delay Condition:				
Two Weeks	4	6	6	3
Four Weeks	7	7	3	4
Eight Weeks	8	7	5	5

Results

Two experimental hypotheses were tested in an analysis of variance.³

H1: The corrective ads would have an immediate negative effect on brand beliefs about the prevention/reduction of colds and sore throats. However, this immediate effect would dissipate over the eight week test period in a manner graphed in Figure 1.

H2: The positive-only copy (non-corrective) condition would have an immediate positive effect on brand beliefs. This immediate effect would decrease over the eight week post period (in a fashion directly opposite to the corrective ads).

To support the first hypothesis of initial lowering of beliefs by the corrective ads and a gradual increase over the eight week period, the order of measurement by length of delay interaction term of the ANOVA should have been statistically significant for attributes relating to colds and cold symptoms prevention. A significant main effect of length of delay by itself would not be a sufficient test of hypothesis one since differences among subjects were expected only when measured after a delayed period. There was no reason to expect differences among these groups when they were each measured immediately after exposure to the corrective ad. The second hypothesis concerning the positive (non-corrective) ad would be supported by a significant main effect of the ad term and a significant three-way interaction of ad and length of delay with order of measurement. The ad copy main effect term would test whether the positive and negative (corrective) ads had, as hypothesized, opposite effects, and the third-order interaction term would test whether these opposite effects had opposite decay effects over time.

Table 3 presents the results of the ANOVA for the sixteen measures. In addition, the analysis of the sum of the products of the importance and beliefs of each of the eight dimensions is shown. For convenience of presenta-

³For the description and results of several hypotheses about the immediate effects of the three corrective advertisements, see Semenik (1976).

tion, only the statistical significance levels of .10 or lower are reported. Two separate ANOVA are reported; the first figure in each column is the results of an ANOVA with four levels of ad copy, and the second figure (in parentheses) is the results of a second ANOVA in which only two levels of ad copy are considered - corrective (negative only or two-sided) and non-corrective (positive-only).

A look at the results reveals several interesting findings not directly salient to the research hypotheses. First, the second ANOVA with ad copy collapsed into two levels seems more statistically powerful. This is probably due to the fact that the ad copy factor in both ANOVAs had very little effect and there was a smaller error term in the second analysis. Second, it seems that being exposed to the deceptive ad shortly before one of the four versions of the test Listerine ad consistently had an effect on both the importance and belief perception on the various attributes (see the column entitled "Exposure" in the table). A check of cell means revealed that the effect of such exposure was always to decrease importance or to decrease the perceived probability that the brand possessed the attribute in question. It appeared that this condition artificially increased the impact of the subsequent corrective ad. Perhaps the first ad sensitized subjects to pay more attention to the later exposure.

As stated earlier, the ANOVA term relevant to the hypotheses about the carryover effects of the corrective ad was the length of delay by order of measurement interaction. Table 3 shows that there was no support for this hypothesis; in no instance was this interaction term statistically significant at the .05 level. The main effect of time was significant for several attributes. However, the absence of a significant interaction term testifies to the fact that, in no case, was the predicted increase in importance or belief probability significantly greater in the delayed conditions (where the increases were predicted) than in the immediate conditions (where there was no reason to expect differences among groups who differed only in that, subsequently, they were to be measured after different delays). For most attributes, there were equal increases for those in the immediate measure condition and those in the delayed condition.

The only result that resembled the hypothesized effects was the importance of "long lasting" which was not explicitly addressed in any of the rested Listerine ads. The time of delay by measurement order interacting term was nearly significant in the first ANOVA ($F = 2.88, 2$ and 98 df, $p = .07$). Table 4 presents the cell means which show no changes in ratings in the immediate conditions and an increase in the delayed conditions. This was the result that was predicted and not found for beliefs relating to the prevention of colds and cold symptoms addressed in the corrective advertising copy. As an example typical of most of the other ratings, Table 5 presents the mean ratings of the target belief "prevents colds/sore throats". It can be seen that an overall increase was found ($F=3.17, 2$ and 110 df, $p=.05$) but very little difference was found between the three immediately measured groups and these same three groups measured, respectively, two, four and eight weeks later.

Given the lack of overall support for the predictions about the decay of beliefs, the question arises as to whether the ads had any effect at all. As previously noted, the ad copy variable showed no significant effects. To test the immediate effects of the ad copy, the positive only (non-corrective) and negative only (corrective) copy appeals were compared to the immediate control group with Dunnett's Comparison test (see Kirk, 1968). These two conditions were chosen because of the conceptual non-ambiguity of the direction of their likely effects.

Table 3

SUMMARY OF ANALYSES OF VARIANCE OF IMPORTANCE OF EIGHT DIMENSIONS AND LISTERINE BELIEFS ON THOSE DIMENSIONS

	Measurement Order	Length of Delay	Exposure to Positive Ad	Ad Copy	Order By Delay
Evaluation of:	df: (1,98)/(1,110) ^a	(2,98)/(2,110)	(1,98)/(1,110)	(3,98)/(1,110)	(2,98)/(2,110)
Appealing taste				(.09)	
Kill germs		(.02)			
Long lasting					.67
Prevents colds/sore throats		.01 (.01)			
Relieves nasal congestion		.07			
Reduce severity of cold/sore throat		.02 (.01)	(.05)		
Reduces coughs/sneezing		.09 (.07)			
Prevents bad breath		(.04)	.05 (.17)		
Beliefs about Listerine:					
Appealing taste	.01 (.01)				
Kills germs		(.04)			
Long lasting			.01 (.01)	.05 (.09)	
Prevents colds/sore throats	.03	(.05)	.05 (.01)		
Relieves nasal congestion			.01 (.01)		
Reduce severity of cold/sore throat	.04		.01 (.01)		
Reduces coughs/sneezing			(.05)		
Prevents bad breath			.05		
$\sum_{i=1}^8 I_i B_i$.01 (.01)	(.04)		

^aIn each main column, the first significant sum and degrees of freedom are for the ANOVA with four levels of the ad copy factor and the second column details the ANOVA with only two levels of the ad factor.

Table 4

MEAN IMPORTANCE OF "LONG LASTING" AS A FUNCTION OF ORDER AND TIME OF DELAY OF MEASUREMENT^a

Group	Order	
	Immediate Post	Delayed Post
Two Weeks Delay	1.91 ^b	2.59
Four Weeks Delay	1.98	1.96
Eight Weeks Delay	1.65	1.35
Immediate Control	1.15	

^aThe lower the number, the more important.

^bFor all three groups, the immediate post measurement followed exposure to the ad. The delayed post followed by the number of weeks indicated in the group label.

Table 5

MEAN RATINGS OF "PREVENTS COLDS/SORE THROATS" AS A FUNCTION OF ORDER AND TIME OF DELAY OF MEASUREMENT^a

Group	Order	
	Immediate Post	Delayed Post
Two Weeks Delay	4.94	4.84
Four Weeks Delay	4.65	4.64
Eight Weeks Delay	4.13	3.86
Immediate Control	4.35	

^aThe lower the number, the higher the rating.

Table 6 shows only a limited number of immediate effects. There was little difference between the positive and negative copy. The main area of immediate impact was the effect of the negative, corrective ad on the importance of four attributes - "appealing taste" and "long lasting" (decrease) and "relieves nasal congestion" and "reduces coughing/sneezing" (increase). Because, as noted earlier, control groups were unfortunately not remeasured in the delayed conditions, Table 6 presents comparisons of de-

layed measures with the immediate control group. Although the validity of these comparisons is tenuous at best, it is interesting to note that, in comparison to the immediate control group, both the importance and Listerine belief on the dimension of "prevents colds/sore throats" went from significantly negative two weeks after the negative copy (corrective ad) exposure to significantly positive at eight weeks. A similar reversal can be noted for the importance of "reduces severity of colds/sore throat" and "prevents bad breath".

Summary and Discussion

The goals of this experiment were to test, for the first time, the carryover effects of corrective advertising and to add to the external validity of past research about corrective advertising. Unfortunately, the results are accompanied by more questions than answers. Although the tested television commercials were explicitly worded, they showed little immediate effects on target beliefs. This non-effect contrasts with the effects found by others including Mazis and Akdinson (1976) with the same advertised mouthwash product. Moreover, no differences between corrective and non-corrective (positive only) ads were found. Finally, no support for the hypotheses about the carryover effects of the corrective ads was found.

Of course, the limited immediate effects made the tests of delayed effects somewhat moot. The hypotheses of rapidly decaying immediate effects assumed that the ads would, as in past research, have the desired initial impact. It might be concluded that the changes adding to external validity in this study also served to limit the effectiveness of the corrective advertisements. Some of the cited past research which resulted in immediate effects of corrective advertising did not focus on a product with familiar claims as the target of the corrective advertising, and thus made it easier for the corrective advertising to be effective. However, the Mazis and Akdinson (1976) experiment, which used the same product, similar beliefs measures, and somewhat similar advertising copy (although a different medium) as this study, did find significant immediate effects on beliefs. Thus, a

Table 6

RESULTS OF BULWELT'S COMPARISON OF POSITIVE AND NEGATIVE AD COPY WITH CONTROL GROUP

Importance of:	Positive vs. Control				Negative vs. Control			
	Imm.	2w	4w	8w	Imm.	2w	4w	8w
Appealing taste					-a	-b		-b
Kills germs		-b	-b			-b	-b	
Long lasting		-b			-b	-b	-b	-b
Prevents colds/sore throats		-b			-a			-a
Relieves nasal congestion			+b	+a	+b			-b
Reduces severity of colds/sore throats				-b	-a			-b
Reduces coughing/sneezing	+b	+b	-b	+b	+b			-b
Prevents bad breath		-a			-a			+b
Beliefs About Listerine on:								
Appealing taste								+a
Kills germs		-a	-b		-b	-b		
Long lasting						-b		
Prevents colds/sore throats				+b	-a			+b
Relieves nasal congestion				+a				+b
Reduces severity of colds/sore throats								
Reduces coughing/sneezing	+a		+a					-b
Prevents bad breath		-b						

a = p < .10

b = p < .05

conclusion that a "stiffer test" of corrective advertising in this study limited the effects of corrective advertising does not seem to be a valid excuse for the non-effects in this experiment.

Although it is conceptually appealing that the proper measures of corrective advertising be belief specific, the few observed effects of the tested corrective ads on rated beliefs and importances were neither intuitively nor theoretically sensible. It seemed that the major impact of the negative corrective copy was to affect importances and not the target ratings. Although the idea that consumers' importances are altered by corrective advertising is not surprising (see Mazis and Adkinson, 1976) it is odd that such effects would be in lieu of effects on beliefs.

If the target beliefs about cold prevention had been immediately lowered as hypothesized, then it would seem that the importance of those beliefs would also decrease since that attribute would be perceived as less of a discriminator among competing brands. Instead, the observed results in Table 6 showed an increase in the importance of the cold and cold symptoms prevention beliefs ("relieves nasal congestion" and "reduces coughing/sneezing") and a decreased importance of (what, conceptually, should be more important) other beliefs ("appealing taste" and "long lasting").

The fact that parallel differences were found in both the immediate and delayed measures is simply baffling. A demand artifact explanation makes no sense since subjects were randomly assigned to a delay condition after they had been immediately measured.

The fairly consistent patterns over two, four and eight weeks suggest the strong possibility that some uncontrolled external force influenced subjects' ratings. It may be that what was thought to be a major strength of the research design - the use of a naturally occurring advertising campaign - proved to be a major problem. Although the mouthwash product did not use the (allegedly deceptive) "colds" campaign during May and June of 1976 when the data were collected, other positive advertising for the product was being telecast. Perhaps this con-

tinued advertising increased all ratings over the period of time in question. If so, the observed increases may not represent, as originally hypothesized, a reversion to former beliefs, but may rather represent the positive effects of Listerine's spring advertising campaign. Alternative speculation would suggest a "ceiling effect" to explain the lack of immediate effects. It could be that the FTC publicity and other available information about the Listerine case in combination with the fact that Listerine had recently ceased its "colds campaign" had already, to some degree, lowered the target beliefs of the corrective advertising. If so, merely one exposure of the corrective ads could not be expected to further decrease beliefs very much. Perhaps the observed increase over time was a decay of the effects of the negative publicity surrounding the case and not, as originally hypothesized, the single corrective ad exposure. Of course, the regrettable lack of delayed control groups becomes a crucial void because of the corresponding lack of ability to answer questions about the effects of any brand advertising during the experimental period. It is also possible that the slightly lower return rate at eight weeks delay represented a lower number of negatively oriented subjects.

The authors plan to re-examine the persistence of effects of corrective advertising. Planned improvements include a strengthening of the power of the experimental design by focusing on the negative-only corrective copy condition with no reminder exposure of the deceptive ad and having a greater sample size per cell. Also, repetition of the corrective ad will be experimentally manipulated to (hopefully) increase its initial impact. Inclusion of proper control groups for the delayed measure conditions (which were planned but not implemented in this experiment) should allow a separation of immediate and delayed effects of the experimental ads from ads exposed outside the laboratory in the period following exposure to the test ads. Finally, improved measurement of cognitive structure following the suggestions of Ahtola (1975) will be incorporated. The questions and problems raised in this study deserve answers and renewed attempts at improved solutions. Despite the methodological challenges, attitude research that goes beyond a static perspective is sorely needed in consumer research.

- Ahtola, O.T., "The Vector Model of Preferences: An Alternative to the Fishbein Model." Journal of Marketing Research, 12(1975), p. 52-59.
- Bartlett, F.C., Remembering: A Study in Experimental and Social Psychology. New York: MacMillan, 1932.
- Boyd, H.W., Ray, M.L., and Strong, E.C., "An Attitudinal Framework for Advertising Strategy," Journal of Marketing, 36(1972), p. 27-33.
- Campbell, D.T., "Perspective: Artifact and Control," in Rosenthal, R. and Rosnow, R. (eds.), Artifact in Behavioral Research, (New York: Academic Press), 1969, p. 351-382.
- Dyer, R.F. and Kuehl, P.G., "The 'Corrective Advertising' Remedy of the FTC: An Experimental Evaluation," Journal of Marketing, 38(1974), p. 48-54.
- Fishbein, M. and Ajzen, I., Belief, Attitude, Intention and Behavior, (Reading, Massachusetts: Addison-Wesley), 1975.
- FTC vs. Warner-Lambert, Initial Decision, FTC Docket 8891, November 25, 1974.
- Gruder, C.; Hennigan, K.; Halamaj, J.; and Cook, T., "An Empirical Test of the Absolute Sleeper Effect Derived from the Discounting Cue Hypothesis." Unpublished manuscript, University of Illinois at Chicago Circle, 1974.
- Hunt, H.K., "Effects of Corrective Advertising," Journal of Advertising Research, 13(1973), p. 15-22.
- Jost, A. "Die Assoziationsfestigkeit in ihrer Abhandigkeit von der Verteilung der Wiederholungen." Zeitschrift fur Psychologie und Physiologie der Sinnesorgane, 14(1897), p. 436-72.
- Kassarjian, H.H., Carlson, C.J. and Rosin, P.E., "A Corrective Advertising Study," in M.J. Schlinger (ed.), Advances in Consumer Research, Volume 2 (Chicago: Association for Consumer Research), 1975, p. 631-642.
- Kirk, R.G., Experimental Design: Procedure for the Behavioral Sciences, (Belmont, California: Brooks/Cole), 1968.
- Lutz, R., "An Experimental Investigation of Causal Relations Among Cognitions, Affect, and Behavioral Intention," Journal of Consumer Research, 3(March 1977), 197-208.
- Mazis, M.B. and Adkinson, J.E., "An Experimental Evaluation of a Proposed Corrective Advertising Remedy," Journal of Marketing Research, 12(May, 1976), p. 178-183.
- Miller, N. and Campbell, D., "Recency and Primacy in Persuasion as a Function of the Time of Speeches and Measurement," Journal of Abnormal and Social Psychology, 59(1959), 1-9.
- Olson, J.C., and Dover, P.A., "Cognitive and Behavioral Effects of Deceptive Advertising," Journal of Marketing Research (1978), forthcoming.
- Papageorgis, D., "Bartlett Effect and the Persistence of Induced Opinion Change," Journal of Abnormal and Social Psychology, 67(1963), 61-7.
- Sawyer, A. G., "The Need to Measure Attitudes and Beliefs Over Time: The Case of Deceptive and Corrective Advertising," in Proceedings, 1976 Educators Meeting, (Chicago: American Marketing Association), 1976 p.380-5.
- Sawyer, A.G., "Measurement of the Effects of Corrective Advertising," in Wind, Y. and Greenberg, M. (eds.), Attitude Research Comes to A Head, (Chicago: American Marketing Association), 1977.
- Sawyer, A.G. and Ward, S., "Carryover Effects in Advertising Communication: Evidence and Hypotheses from Behavioral Science," Working Paper No. 76-122, Marketing Science Institute, December, 1976.
- Semenik, R.J., "Message Forms and Effective Corrective Advertising: An Experiment," unpublished doctoral dissertation, The Ohio State University, 1876.
- Weiss, W., "A Sleeper Effect in Opinion Change," Journal of Abnormal and Social Psychology, 48(1953), p. 173-80.
- Wilkes, R.E. and Wilcox, J.B., "Recent FTC Actions: Implication for the Advertising Strategist," Journal of Marketing, 38(1974), p. 55-61.
- Wilkie, W. and Gardner, D., "The Role of Marketing Research in Public Policy Decision Making." Journal of Marketing, 38(1974), 38-47.

Barry R. Schlenker, University of Florida

Abstract

A social identity approach is described which treats attitude expressions as interpersonal communications that present a particular image of the respondent to self and others. The approach predicts when people will associate themselves with particular symbols, and when attitude change will follow pro and counter-attitudinal behaviors. These predictions often differ from those advanced by dissonance and self-perception theories.

The concept of attitude has occupied a central and almost hallowed place in the explanatory systems of theorists interested in human behavior. The field of social psychology was once defined as the study of attitudes (Thomas & Znaniecki, 1918), and the utility of the concept for the field of consumer behavior has hardly gone unrecognized: "No other single psychological construct has permeated consumer research as has the construct of attitudes" (Jacoby, 1976, p. 337).

Attitudes and Social Identity

Attitude is an intrapersonal construct designed to account for different individuals' consistent favorable or unfavorable response patterns vis-a-vis particular social objects and/or events. Attitude is usually defined as the amount of positive or negative affect held toward a particular object and/or event. However, all expressions of attitudes, whether given through gross overt behaviors, verbal statements, or questionnaire responses, are interpersonal communications. Like other actions, attitude expressions present a particular image of the respondent to himself and to others. Attitude responses carry diagnostic value, since they can associate the person with particular groups of people (e.g., liberals vs. conservatives), reveal aspects of the person's personality (e.g., simple vs. complex; stable vs. unstable; optimistic vs. pessimistic), reveal the person's positive or negative orientation toward the social interaction and social setting (e.g., friendly vs. hostile), reveal the person's desired role in the interaction (e.g., dominant vs. submissive), and can be used to gain approval and avoid disapproval (e.g., conformity vs. nonconformity). In short, attitude expressions communicate aspects of the person's social identity and world view and set the tone for interactions. Even "private" attitudes, expressed only to oneself, occur in the context of imaginary social settings and serve the same functions in our private worlds. Every response used to infer our own and others' attitudes (except for physiological measures, which are notoriously unreliable) exists in a matrix of real or imagined social interaction, and can be used to influence the opinions and actions of others.

This focus on the interpersonal nature of attitudes is similar in some respects to the functional approach of the Michigan school (Katz, 1960), which examined attitudes according to their instrumental, ego-defensive, knowledge, and value expressive uses. Ultimately, these functions can be compressed to two: an informational facet, which involves cognitively organizing information about ourselves, others, and the world in a way that can guide actions; and an instrumental, or hedonic facet, which involves acting to maximize desirable outcomes and

minimize undesirable outcomes. Expectancy-value or subjective expected utility models (e.g., Anderson, 1971; Fishbein & Ajzen, 1975; Jones & Gerard, 1967; Rosenberg, 1956) come closest to capturing these facets as they pertain to the conceptualization and measurement of attitudes. However, researchers who have worked with these positions have generally not devoted sufficient attention to the symbolic, interactional, and social identity aspects of attitudes, except indirectly through the assessment of the probability or utility components. Focusing on these aspects, though, generates hypotheses which might otherwise remain hidden and provides for the integration of seemingly disparate phenomena (e.g., Freudian defense mechanisms, reactions to inequity, dissonance) into a common framework. Thus, while accepting the basic usefulness of such models, it becomes worthwhile to explore avenues opened by a focus on social identity and attitudes.

Social Identity and Self-Presentation

The presentation of self is the activity of constructing and maintaining particular identities in social life. Identity is the composite of the social attributes associated with an individual.² Each attribute creates a particular image of the person. Everyone--business person, politician, housewife, student, professor--develops an identity through and for social interaction and projects various aspects of that identity to particular audiences. Through impression management strategies people control how they see themselves and how they expect to be seen and treated by others. People respond to others on the basis of the identities those others create, so everyone finds it to his or her advantage to control the image of self presented for public and private consumption (Goffman, 1959; Jones & Wortman, 1973; Schlenker, 1975a).

The social identity approach derives in part from symbolic interactionism, a model of people which views them as symbol users and interpreters. According to the position, personality develops from and is inseparable from real or imagined social interactions (Carson, 1969). The focus is on the integral relationship between the person, others, society, and social symbols (cf. Brittan, 1973). Goffman's (1959, 1967, 1971) dramaturgical approach to social behavior developed from symbolic interactionism and views people as actors who perform roles available in society, managing the impressions they give to others in order to lay claim to specific social attributes. Although Goffman's work tends to minimize psychological factors and focuses upon interaction rituals, there is no reason to exclude internal, personal factors when studying self-presentation.

²It is worthwhile to distinguish between the self-concept and social identity. The self-concept is defined as the theory an individual has constructed about himself or herself (Epstein, 1973), and consists of perceptions of personal attributes and relationships between them. Social identity refers to "the individual's major role and social-type categorizations" (Gordon, 1968, p. 118). It describes what, who, and where the person is in social terms, and is not simply substitutable for the word self due to the more extensive social evaluation implied.

¹The present paper was facilitated by National Science Foundation grant BNS77-08182.

Constructing Identities: Attitudes and Identification

It is assumed that people manage the impressions they create to project a particular identity, both constructing it and maintaining it against actual or potential threats. Attitude expressions partially define identity and can arouse identity-maintenance concerns. Naturally, attitude topics vary in the degree to which they are relevant to identity concerns; the social judgment theory concept of ego-involvement is one way of conceptualizing such a continuum (Sherif & Hovland, 1961; Sherif et al., 1965).

Products similarly vary in the degree to which they have identity relevance. Some products can capitalize on their identity value; e.g., a Cadillac is a symbol of affluence and status, a Porsche is a symbol of a sporty life style. Other products, which initially might seem to suffer a handicap in this regard for one or another reason, can have symbolic value created for them; e.g., a Volkswagon "bug" seems at first glance to be ugly, but became a symbol of a pragmatic, environmentally-concerned life style. Even the use of products which might not appear to have any relevance for one's identity can be surprisingly revealing, at least to particular people. For example, a mother might never forgive her daughter-in-law for serving her son frozen or canned vegetables rather than fresh ones. To this particular mother, such behavior connotes laziness and indifference to her son's nutritional needs, and is an expression of the daughter-in-law's flawed personal qualities.

It often seems to be the goal of advertisers to increase the positive identity relevance of particular product lines or brands as a means of producing attitude change and ultimately affecting purchasing decisions. Desires for new products can be created by focusing on identity concerns (e.g., unless deodorants are used, doubts are raised about one's viability as a social participant) or previously neutral product brands are associated with existing relevant attributes (e.g., use of "Brut" deodorant identifies a "macho" individual). The classic techniques of testimonials and getting on the band wagon can be seen as attempts to create bonds between the consumer, some desirable image (e.g., attributes of the "star" or qualities of the symbol), and the product. "Join the Pepsi generation" (and have all the fun it entails for those who do); be like Farrah Fawcett-Majors or Joe Namath; smoke the cigarette of the rough-and-ready cowboy. Numerous studies have been devoted to specifying images associated with products (Jacoby, 1976).

From a social identity perspective, such attitude change is exemplary of a more general hypothesis: People act to increase their personal association with desirable attributes and outcomes, and act to decrease their personal association with undesirable attributes and outcomes. Personal association is defined as the probability that a particular attribute describes the individual; or in the case of outcomes, the degree to which the person is seen as responsible for producing the outcome. Desirability is defined as the worth of a particular attribute or outcome to the individual in a particular situation; it is defined idiosyncratically rather than normatively.

Research from a variety of areas provides strong support for the association hypothesis. In attributing the "causes" of particular actions, most people retrospectively report greater responsibility for successful task performances than for failing ones (e.g., Mynatt & Sherman, 1975; Schlenker, 1975b; Schlenker & Miller, 1977; Schlenker et al., 1976; Snyder et al., 1976; Wolosin et al., 1973; Wortman et al., 1973), and this is true whether the person worked on the task alone or with a group of others. People also more readily accept the implications of favorable rather than unfavorable personality information (e.g., Eagly, 1967; Jones, 1973;

Steiner, 1968). Even when a person is objectively non-responsible for producing a desirable outcome, the same proclivities are evidenced. Cialdini et al. (1976) found that more students wore school-identifying products (e.g., school sweatshirts) after their school's football team had been victorious rather than nonvictorious; and that they used the pronoun "we" more when talking about a school victory rather than a defeat. The latter effect was strongest when the students had previously suffered a personal failure that threatened their self-images. When important self-images are threatened or otherwise in doubt, people display a greater association/dissociation effect.

As the above suggests, individuals often prefer to associate themselves with attributes that are evaluated by most people as desirable, e.g., competence, attractiveness. Such findings are frequently explained as being the result of self-esteem needs, social approval needs, or cognitive consistency needs. But it seems somewhat simplistic to propose that people's identity-relevant actions are taken merely to maximize self-esteem and/or to maximize approval from others. It would then be difficult to explain why some people present themselves as incompetent, hostile, unfriendly, unworthy of responsibility for success, or even out of touch with reality and mentally ill (cf. Braginsky et al., 1969; Carson, 1969). Braginsky et al. (1969), for example, found that schizophrenics' self-presentational behaviors were affected by the identity they wished to project to particular audiences. Mental patients who were old-timers in an institution and who enjoyed open ward privileges were asked to complete mental tests or talk with a psychiatrist. These patients were assumed to prefer the freedom, with attention from the staff, provided by the open wards, and would want to avoid being seen as either healthy enough to be sent home or ill enough to be confined to restrictive closed wards. When the patients believed that the test/interview was for the purpose of diagnosing their mental health, with high scores indicating they should be sent home, they presented themselves as mentally ill. They gave frequent reports of hallucinations and showed other signs of abnormal, incompetent functioning. But when they believed that the test/interview was for the purpose of diagnosing their mental illness, with high scores indicating they should be placed in undesirable closed wards, they presented themselves as mentally healthy. They gave few reports of hallucinations and showed signs of normal, competent functioning. Thus, schizophrenics, often defined as being out of touch with reality and incapable of normal functioning, were quite adept at modifying their self-presentations to fulfill their interaction goals.

Similarly, identity-relevant actions don't seem motivated by "needs" for cognitive consistency. If they were, it would be difficult to explain the prevalence of inconsistent belief systems or belief-behavior relationships (cf. Tedeschi et al., 1971). Schlenker (1975a), for example, found that people who view themselves as "failures" will present themselves in a self-effacing manner when the failures are publicly known but will present themselves quite positively when the failures are unknown to the audience.

Such data falls into place when viewed from a social identity perspective. The desirability of an attribute or outcome can only be assessed with reference to the identity the individual is attempting to project. As Braginsky et al., (1969) demonstrated, different attributes acquire different values depending upon one's goals in a situation. Further, certain attributes can only be claimed by certain people. Most people might like to be seen as intelligent, competent, witty, attractive, respected, and so forth. But only certain individuals can successfully lay claim to such attributes; reality does impose constraints on projected identities. Inaccurate claims to social attributes produce

disruption of interactions, anxiety, guilt, and the possible censure and punishment of the offender (Goffman, 1959, 1967). Not only does an illicit identity (i.e., one the individual has no "right" to claim) produce embarrassment and anxiety for the discovered offender, it creates similar states for those who are duped into believing the misrepresentation and then realize their mistake. Schlenker (1975a) supported the hypothesis that people must believe they can successfully (without irrefutable challenge) act in accord with the identities they project before they will present themselves in a normatively positive fashion to an audience.

Claimed images or attributes can create obligations for the person. An obligation is defined as the belief that certain physical characteristics or behaviors should accompany claims to a particular image or attribute. A person who claims intellectual prowess should be able to demonstrate it; a person who claims physical beauty should possess some such recognizable features. Attributes (and products) vary in the degree to which they create obligations, and people vary in the degree to which they perceive obligations. A man might avoid a "Joe Namath" cologne if he believes wearing it would make him the object of jests about his baldness, pot belly, lack of coordination, and failures with women. A second man with a similar set of characteristics might not believe that any obligations derive from the use of the product (and thus not expect such chiding), and revel in his fantasies of himself as playboy extraordinaire.

Images or attributes also vary in the degree to which claiming them produces public (vs. only private) concerns. The personal desirability of a normatively desirable attribute will be low when the identity it implies creates public obligations which the person believes he or she cannot satisfy. But if a normatively desirable attribute only creates private obligations (e.g., those of a Walter Mitty fantasy world) and no public ones, its personal desirability will be high. The use of some products (e.g., "personal" items meant to be concealed), even if it would create obligations if publicly known, can be hidden away from public scrutiny.

In the realm of consumer behavior, this translates into the prescription that one must be careful of the type of image associated with particular products. The individual must attach a high personal desirability to the image. But this will only occur when he or she favorably evaluates the particular image in others, and either (a) believes that there are no public obligations associated with the image, or (b) believes that he or she can fulfill whatever public obligations are perceived, or (c) believes that although there might be public obligations that can't be fulfilled, he or she can "hide" public use of the product and not have to fulfill them. If public obligations exist which cannot be hidden and cannot be personally fulfilled, normatively desirable images will be rejected.

This would seem to be a particularly fruitful area for consumer research. Although worthwhile data do exist on product and brand images, to my knowledge nothing has been done with respect to (a) obligations that derive from those images, (b) consumers' abilities to fulfill whatever obligations exist, and (c) the public vs. privateness of whatever obligations may exist. Knowledge of the interrelationship between these factors relative to particular products could provide a more adequate predictor of purchasing decisions and behaviors.

Maintaining Identities: Attitudes, Accounting, and Counterattitudinal Actions

One of the more salient areas of attitude research has been on the relationship between behaviors and subsequently expressed attitudes, particularly with respect

to counterattitudinal behavior. Attitude change following counterattitudinal behavior has typically been explained via dissonance theory (Aronson, 1968; Festinger, 1957) or self-perception theory (Bem, 1967, 1972). The social identity approach offers an alternative conceptualization of the area which focuses on the interpersonal concerns raised by such actions (Riess & Schlenker, 1977; Schlenker, 1973, 1975c; Schlenker & Forsyth, 1977; Schlenker & Schlenker, 1975).

The counterattitudinal advocacy paradigms employed by psychologists typically induce subjects to lie, cheat, harm others, refrain from doing what they want to do, or otherwise make themselves appear to be immoral, unattractive, incompetent, or irrational. In other words, they place subjects in a predicament that could threaten their projected identities. Following actions which create a predicament (i.e., ones which are embarrassing, undesirable, illegal, anti-normative, or disruptive) people try to account for their behaviors. Accounting is an attempt to describe (interpret) the action and its consequences in personally desirable terms. Typically, this translates as interpretations which are socially acceptable (Lyman & Scott, 1968). People are "scientists" in the sense that they try to understand and explain the causes of events through attributions. But accounting can go beyond "logical" attribution in that a search is made to "explain" one's otherwise undesirable actions in desirable ways. Accounting involves the use of attributions and attitude expressions in the service of maintaining a particular identity.

Although there are numerous accounting tactics, two general classes of responses predominate--excuses and justifications. Excuses are attempts to remove from oneself personal responsibility for the predicament. Excuses include tactics such as (a) denying one did it, (b) stating it was an accident or the consequences were unforeseeable, (c) citing external coercive pressures (e.g., "He made me do it"), and (d) citing internal coercive pressures (e.g., "I couldn't help it, I was [starving] [drunk] [mentally ill] [possessed by demons]"). Justifications allow a person to admit responsibility for the predicament but alter the perceived negativity of the consequences of the predicament. For example, one might justify harming another by noting why the other person deserved it (and hence the act was not "bad") or by minimizing the amount of harm done ("It was only a scratch"). Thus, successful excuses allow one to escape from the predicament by minimizing personal responsibility, while successful justifications allow one to escape by minimizing the aversive consequences.

It is hypothesized that the particular type of account which is used in any situation will depend upon (a) its subjective probability of extricating the person from the predicament or minimizing the potential punishment, and (b) the utility of the perceived outcomes it generates (the degree to which it eliminates or minimizes the potential negative sanctions from observers and from self guilt). People will use that accounting tactic with the most positive (or least negative) subjective expected utility. The characteristics of the actor, situation, and audience affect these probabilities and utilities, since they affect the standards involved in judging the action and the degree to which a particular account could be challenged or refuted by the "facts." As this implies, accounts which can be obviously refuted by known information are least likely to be used because they are easiest to challenge, making the person not only still accountable for the original action but also sanctionable for the attempted deceit. Space does not permit a full exploration of accounting here, but two points are worth noting. First, predictions can be made about when particular types of accounting tactics will be used. Second, the accounting tactics mentioned above bear a striking resemblance to phenomena previously discussed under the headings of (a) modes of dissonance

reduction, (b) equity restoration techniques, and (c) defense mechanisms in psychopathology.

People frequently internalize the accounts they deliver and come to believe their own interpretation of the situation, producing relatively enduring attitude change or restructuring. It is hypothesized that people will internalize their accounting tactics when they negatively evaluate the action that created the predicament. If a person does not negatively evaluate the act, or feels the act is much less negative than does an audience, an account might be used to placate the audience but will be recognized by the actor as a "necessary lie" (cf. Schlenker & Forsyth, 1977, for an elaboration of this hypothesis). When "necessary lies" are told, it can complicate the person's existence; he or she must then keep multiple sets of books to recall which accounts were delivered to which audiences. Thus, an account which is accepted by both self and others facilitates one's effective functioning.

The amount of accounting necessary following a predicament is directly related to both personal responsibility for the predicament and the magnitude of its negative consequences. Responsibility is defined as the attribution that an individual freely and purposefully behaved to produce the primary consequences associated with the action, and did not engage in the behavior because of coercive pressures, accident, or to achieve other associated consequences. The more negative the action and the greater the responsibility of the actor, the more imperative it is to eliminate, or at least minimize, the potential guilt and punishment.

Consistent with the accounting hypothesis, it has been found that significant attitude change occurs following counterattitudinal actions only when a person appears to be personally responsible for engaging in the behavior and the behavior produces aversive consequences (Calder et al., 1973; Collins & Hoyt, 1972; Hoyt et al., 1972). Envision the typical paradigm. A subject is given either high or low choice (high vs. low responsibility, respectively) about making a speech which persuades or does not persuade (high vs. low harm, respectively) an audience. If no choice exists, the person has a ready-made visible excuse for the behavior and does not have to further account for it. If no harm is done, there is no need to account for the action. Only when responsibility and aversive consequences coexist is accounting necessary. But how can a high choice, high harm subject account for the behavior? The action can't be excused, since it was emphasized that the choice to perform it was solely his. It can be justified, though, by showing why little or no harm occurred. If the subject expresses attitudes which are roughly congruous with the position taken in the speech, then no real harm has occurred--after all, the subject personally endorses the beliefs the audience adopted. The counterattitudinal behavior becomes neither a lie nor harmful to others.

Research examining social identity theory indicates that responsibility and consequences are important not because they impact on cognitive consistency, but because they define the nature of the social identity dilemma (e.g., Forsyth et al., 1977; Riess & Schlenker, 1977; Schlenker, 1975a, 1975c; Schlenker & Forsyth, 1977; Schlenker & Schlenker, 1975). In many (though not all) counterattitudinal behavior situations, social identity theory makes the same predictions as does Aronson's (1968) revision of dissonance theory, which stresses that dissonance is produced when the implication of a behavior is inconsistent with one's self-concept. However, the two approaches differ in the hypothesized mediating mechanisms behind attitude change. Dissonance theory focuses on cognitive consistency-inconsistency, while social identity theory focuses on accounting following social predicaments. The difference is best illustrated by the fact that most counterattitudinal behaviors do two

things: (1) produce a predicament for the person since he or she has, by definition, lied and perhaps harmed others, thus being subject to sanctions and (2) produce cognitive inconsistency. This natural covariation has perhaps caused dissonance theorists to focus on the wrong variable. According to dissonance theory, motivation arises from cognitive inconsistency; according to social identity theory, motivation arises from the nature of the identity predicament. Although these variables covary in most situations, they can be separated such that a "lie" is made to produce positive rather than aversive consequences. When this is done, data indicate that inconsistency *per se* is particularly non-troublesome, and people behave according to predictions derived from social identity theory rather than dissonance or self-perception theory (Schlenker & Schlenker, 1975).

It is necessary to elaborate some of the implications of the above. The above does not imply that people do not organize attitudes and beliefs in a way that frequently gives the appearance of logical (or psychological) consistency. They do. All models of information integration (e.g., Anderson, 1971; Osgood & Tannenbaum, 1955; Wyer, 1974) are in some sense "consistency" models, since they describe how information is processed to achieve a "consistent" behavioral orientation. However, information integration models do not automatically lead to the conclusion that inconsistency has major motivational by-products that steer social behaviors in a particular direction. Only the so-called cognitive consistency theories (balance theory, dissonance theory) posit such drive-reduction needs arising from cognitive inconsistency. It is one thing to discuss the process of information integration, another to assume that cognitive inconsistency *per se* produces psychological needs which must be satisfied. The fact that people do weight and integrate information in order to make sense out of their environment and to function effectively in it is certainly compatible with (indeed, assumed by) social identity theory. Similarly, "incentive theory" (Elms, 1967; Janis & King, 1954; King & Janis, 1956) ideas about the effects of cognitive contact and biased scanning on attitude change following role-playing are quite compatible with information processing models. However, most research attention has been devoted to people's attitudinal reactions to predicaments. In such situations, motivation does not appear to arise from intrapsychic, cognitive consistency pressures, but from the nature of the predicament for social identity.

Social identity research suggests that many of the dissonance-type ideas about counterattitudinal behavior, effort justification, and so forth, might not have the marketing implications previously thought. For example, getting a consumer to reach to a higher shelf to get a product (exerting maximal effort) might not automatically lead to more favorable attitudes toward the product. It might do so only to the extent that the behavior created a predicament for the individual. Say, he was asked by another shopper why he strained to get the product when an equivalent one was in easy reach, believed so doing was somewhat irrational, believed it would harm his identity if he admitted this was irrational, and believed that an account based on product-liking would be most effective. It would be rare for a person to encounter this specific sequence of events. Or, instead of using "free giveaways" of new products, dissonance theorists might advocate making consumers pay some small amount for it or making them mail a postcard (exerting effort) rather than having it left on their doorstep. Once again, unless a predicament is generated, the strategy does not seem likely to have the most positive effects. Hedonic considerations suggest fewer people will be exposed to the products, and those that are will probably not develop any more favorable attitudes than otherwise. Marketing strategies based on "insufficient justification" do not appear to have strong data

to recommend them.

The only marketing area that does seem to yield data consistent with the "insufficient justification" idea involves the use of pricing techniques. Some dissonance researchers have recommended that products be sold for the largest amount possible (that will still get purchases) so that dissonance is aroused (Doob et al., 1969). Yet when a direct relationship is found between pricing and liking for a product, it is probably best to conceptualize it as due to economic evaluation processes; the relationship between price and perceived product quality has generated more laboratory research during the past eight years than any other topic in consumer behavior (Jacoby, 1976). Many products are simply perceived as more valuable the higher their price, at least up to some upper limit (Gabor & Granger, 1966). This is true irrespective of whether a purchase has been made.³

Maintaining Identities: Attitudes, Acclaiming, and Proattitudinal Actions

The social identity approach is built around the self-presentational nature of attitude expressions; it focuses upon the potential consequences of attitude expressions and the ways the consequences and the situation affect attributions which can be made about the person by himself and others. As was seen, many counterattitudinal actions produce consequences which are identity-threatening and which require accounting tactics. A complimentary process should occur following proattitudinal actions. Proattitudinal actions are often identity-supporting, and, depending upon the nature of the action, might produce consequences which are highly desirable. Hence, the association hypothesis indicates that people should want to appear personally responsible for relevant proattitudinal behaviors. For example, a person who contributed to a charity would prefer the attribution that the donation was made because he believed in the cause, and would want to avoid attributions such as "He did it because it is expected of those in his societal position," or "He did it because he was pressured into it by the canvasser for the charity." Responsibility implies that the person is entitled to the appropriate rewards, credit, respect, and feelings of self-worth that derive from the action. As the desirability of the consequences increases, so should the value of the accompanying image and personal association.

If an individual already appears associated with the behavior and its obviously desirable consequences, no further actions are necessary; he or she can reap the appropriate rewards. However, if the actor does not appear to be responsible for the action or if the consequences appear to be less desirable than the actor believes they should appear, acclaiming tactics will occur. Acclaiming is an attempt to describe (interpret) an action and its consequences in ways which reflect favorably on the actor. The two most salient acclaiming tactics are entitlements, which involve attempts to increase one's responsibility for an action and its consequences, and enhancements, which involve attempts to increase the apparent positivity of the consequences of the action. As with accounting, predictions can be derived about

³This is not to suggest that postdecisional distortions of the values of accepted and rejected alternatives do not occur. However, this well documented phenomenon does not follow exclusively from dissonance theory. Virtually identical predictions can be derived from conflict models, the idea of unequivocal behavior orientation (Jones & Gerard, 1967), social identity theory (based on derivations of the association hypothesis), and self-perception theory. Though the mediating processes may be debated, the end results relevant for consumer behavior seem to be comparable for each approach.

which acclaiming tactics will be used in particular situations based on the tactics' subjective expected utilities.

It is hypothesized that when environmental variables decrease the amount of responsibility that can be attributed to a person for proattitudinal actions, the more desirable the actions and their consequences, the more the person will use acclaiming tactics. Responsibility can be demonstrated by increasing the extremity of one's position on the topic or by appearing highly involved with the topic, such as by stressing its personal importance, its value to society, and one's behavioral commitment to it. Extremity and involvement are typically conceptualized as dimensions of one's attitude toward an issue (Scott, 1968). When a person already appears responsible for a desirable action, little or no additional acclaiming is required; responsibility and recognition are already theirs. But when environmental pressures decrease responsibility and dissociate the person from the positive consequences, he or she should use acclaiming tactics to increase the likelihood that responsibility is attributed.

As an example of the application of this hypothesis in role-playing situations, consider a politician who delivers a speech which coincidentally favors the popular side of some local issue. Naturally, the politician would prefer the audience to make the attribution that the speech reflected his personal attitudes (personal attribution) rather than that the speech merely reflected the desire for reelection (role attribution). To insure that the origin of the speech is attributed to personal attitudes, the politician might demonstrate increased gesticulation, hand-pounding, and voice inflection during the speech (as compared to a situation in which responsibility was not in doubt). In addition, he should evidence more favorable attitudes toward the issue at the conclusion of the speech than were held by him the previous day, at least up to some maximal point. This show of involvement with the issue would serve as a means of convincing the audience, and perhaps himself, that the behavior was desirable and was due to personal feelings rather than external role constraints.

Consistent with the above, research has found that when important positive consequences are produced by behaviors (e.g., delivering a proattitudinal speech), subjects subsequently express more favorable attitudes toward the topic when they are given no choice rather than high choice about engaging in the behavior; the effect has been found following both proattitudinal (Schlenker & Riess, 1977) and counterattitudinal (Schlenker & Schlenker, 1975) actions. These results fail to support predictions of the opposite effect (greater attitude change following high rather than low decision freedom) derived from self-perception theory.

These findings indicate that when the consequences of an action are identity-relevant and important, the type of "insufficient justification effect" which occurs will depend upon whether the consequences are desirable or undesirable. When the consequences are undesirable, insufficient justification (i.e., high responsibility) for the behavior is directly related to subsequent attitude change; as choice increases, so does attitude change. When the consequences are desirable, though, insufficient justification is inversely related to subsequent attitude change; as choice increases, attitude change decreases.

These results do not imply that self-perception principles are majorly incorrect. Self-perception theory, in conjunction with even more general ideas about attribution, has been quite successful in predicting attitudes in situations where (a) the consequences of the action are minimal and social identity concerns are not aroused (e.g., Taylor, 1975), and (b) subjects are given

false or misleading information about their internal states and must make sense out of the information available to them (e.g., Nisbett & Valins, 1972). As the desirability of an action increases, though, social identity concerns become engaged and people appear to follow a pattern of rationalization rather than "rational" information processing.

Conclusions

The social identity approach highlighted here has potential for integrating a wide variety of phenomena under a common rubric. The approach weaves together symbolic interaction (people are symbol users and interpreters), attribution theory (behaviors have social meanings which reflect on the identity of the actor with certain probabilities and consequences), hedonic principles (people act to maximize subjective expected utility), and social influence (people use impression management strategies for accounting and acclaiming purposes, which influence self and others).

In the realm of attitudes, the approach is not competitive with information processing models or with functional approaches to attitudes. Rather, it complements these by focusing on the communicative, interpersonal, social, and symbolic uses of attitude expressions. It emphasizes both rationalization and rational information processing. Rationalization is not conceptualized as an aspect of Freudian intrapsychic conflicts determined sometime during the first five years of life, or as due to needs for cognitive consistency. Instead, it is viewed as an identity protecting and enhancing tactic that arises from the nature of the social interaction process. People are "action centers" (Heider, 1958) who intend to produce certain effects and who are held accountable for what they do. The rewards and punishments (both from others and from pride or guilt) associated with accountability provide a powerful hedonic impetus for rationalization.

At one level, it has been argued that the basic motivational processes described by dissonance theory may well be in error--needs for cognitive consistency do not appear to have the omnipotent driving force that has been attributed to them. At another level, the kinds of predictions typically made by dissonance theory can be reconceptualized as a "special case" of the more general social identity approach. Such a possibility has been indirectly suggested by Zimbardo (1969, p. 15), who believes that Festinger's theories of social communication, social comparison, and dissonance, "may be subservient to more basic phenomena which characterizes this approach as a 'face-saving theory,' in which the individual is motivated to modify and distort both internal and external reality in order to make them appear consistent with having made the 'correct' decision."

Although the attribution process is salient in the social identity approach, attributions serve as the foundation for information about one's identity. These attributions can be biased by one's self-interests in order to control identity; attributions are acted on and manipulated through accounting and acclaiming tactics. From this perspective, attribution does not play the passive role it does in self-perception theory, where attributions about our attitudes are the end points of the attitude change process. There is rarely, if ever, only one attribution that can be made about an action. People exercise the opportunity to sort through the attributional possibilities until one or more can be found which best serve self-interests and establish the preferred type of social identity.

Finally, the social identity approach is most directly applicable to symbols that are (or can be made) at least somewhat identity relevant. The boundaries of identity relevance are cloudy at the moment, since relatively

little empirical work has been directed at specifying when people evaluate a symbol as relevant versus irrelevant to their identity. But it seems reasonable to suggest that many products have some identity relevance, or at least that many products can take on such relevance through particular types of advertising campaigns. Campbell's soups, for example, once stressed their relevance to the image of the stay-at-home housewife by emphasizing attributes associated with mother's good home cooking. Today, changing times have dictated another strategy; the value of the product is stressed for the image of liberated women who want to provide proper nourishment for their families while escaping the house quickly for an afternoon game of tennis. The degree of identity relevance of particular products or product brands is an empirical question, but it would seem that a large number of products already have the requisite relevance to make social identity hypotheses applicable.

References

- E. Aronson, "Dissonance Theory: Progress and Problems," In R. Abelson, E. Aronson, W. McGuire, T. Newcomb, M. Rosenberg and P. Tannenbaum (Eds.), Theories of Cognitive Consistency: A Sourcebook, (Chicago: Rand McNally, 1968)
- N. H. Anderson, "Integration Theory and Attitude Change," Psychological Review, 78 (1971), 171-206.
- D. J. Bem, "Self-perception: An Alternative Interpretation of Cognitive Dissonance Phenomena," Psychological Review, 74 (1967), 183-200.
- D. J. Bem, "Self-perception Theory," In L. Berkowitz (Ed.), Advances in Experimental Social Psychology (New York: Academic Press, 1972)
- B. M. Braginsky, D. D. Braginsky and K. Ring, Methods of Madness: The Mental Hospital as a Last Resort (New York: Holt, Rinehart and Winston, 1969)
- A. Brittan, Meanings and Situations (Boston: Routledge and Kegan Paul, 1973)
- B. J. Calder, J. Ross and C. A. Insko, "Attitude Change and Attitude Attribution: Effects of Incentive, Choice, and Consequences," Journal of Personality and Social Psychology, 25 (1973), 84-89.
- R. C. Carson, Interaction Concepts of Personality (Chicago: Aldine, 1969)
- R. B. Cialdini, R. J. Borden, A. Thorne, M. R. Walker, S. Freeman and L. R. Sloan, "Basking in Reflected Glory: Three (Football) Field Studies," Journal of Personality and Social Psychology, 34 (1976), 366-375.
- B. E. Collins and M. F. Hoyt, "Personal Responsibility-For-Consequences: An Integration and Extension of the 'Forced Compliance' Literature," Journal of Experimental Social Psychology, 8 (1972), 558-593.
- A. N. Doob, J. M. Carlsmith, J. L. Freedman, T. K. Landauer and S. Tom, Jr., "Effect of Initial Selling Price on Subsequent Sales," Journal of Personality and Social Psychology, 11(1969), 345-350.
- A. H. Eagly, "Involvement as a Determinant of Responses to Favorable and Unfavorable Information," Journal of Personality and Social Psychology, 7(1967), Whole No. 643.
- A. C. Elms, "Role-Playing, Incentive and Dissonance," Psychological Bulletin, 68(1967), 132-148.

- S. Epstein, "The Self-Concept Revisited: Or a Theory of a Theory," American Psychologist, 28(1973), 404-416.
- L. Festinger, A Theory of Cognitive Dissonance (Evanston, Ill.: Row, Peterson, 1957)
- M. Fishbein and I. Ajzen, Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research (Reading, Mass.: Addison-Wesley, 1975)
- D. R. Forsyth, M. Riess and B. R. Schlenker, "Impression Management Concerns Governing Reactions to a Faculty Decision," Representative Research in Social Psychology (1977), in press.
- A. Gabor and C. W. J. Granger, "Price as an Indicator of Quality: Report on an Enquiry," Economica, 33(1966), 43-47.
- E. Goffman, The Presentation of Self in Everyday Life (Garden City, N.Y.: Doubleday Anchor, 1959)
- E. Goffman, Interaction Ritual: Essays on Face-to-Face Behavior (Garden City, N.Y.: Doubleday Anchor, 1967)
- E. Goffman, Relations in Public (New York: Basic Books, 1971)
- C. Gordon, "Self-conceptions: Configurations of Content," In C. Gordon and K. J. Gergen (Eds.), The Self in Social Interaction (New York: Wiley, 1968)
- F. Heider, The Psychology of Interpersonal Relations (New York: Wiley, 1958)
- M. F. Hoyt, M. D. Henley and B. E. Collins, "Studies in Forced Compliance: The Confluence of Choice and Consequences on Attitude Change," Journal of Personality and Social Psychology, 22(1972), 1-7.
- J. Jacoby, "Consumer Psychology: An Octennium," In M. R. Rosenzweig and L. W. Porter (Eds.), Annual Review of Psychology, 27(Palo Alto, Ca.: Annual Reviews, Inc., 1976)
- I. L. Janis and B. T. King, "The Influence of Role-Playing on Opinion Change," Journal of Abnormal and Social Psychology, 49(1954), 211-218.
- E. E. Jones and H. B. Gerard, Foundations of Social Psychology (New York: Wiley, 1967)
- E. E. Jones and C. Wortman, Ingratiation: An Attributional Approach (Morristown, N.J.: General Learning Press, 1973)
- S. C. Jones, "Self and Interpersonal Evaluations: Esteem Theories Versus Consistency Theories," Psychological Bulletin, 79(1973), 185-199.
- D. Katz, "The Functional Approach to the Study of Attitudes," Public Opinion Quarterly, 24(1960), 163-204.
- B. T. King and I. L. Janis, "Comparison of the Effectiveness of Improvised Versus Non-Improvised Role-Playing in Producing Opinion Changes," Human Relations, 9 (1956), 177-186.
- S. M. Lyman and M. B. Scott, A Sociology of the Absurd (New York: Appleton Century-Crofts, 1970)
- C. Mynatt and S. J. Sherman, "Responsibility Attribution in Groups and Individuals: A Direct Test of the Diffusion of Responsibility Hypothesis," Journal of Personality and Social Psychology, 32(1975), 1111-1118.
- R. E. Nisbett and S. Valins, "Perceiving the Causes of One's Own Behavior," In E. E. Jones, D. E. Kanouse, H.H. Kelley, R. E. Nisbett, S. Valins and B. Weiner (Eds.), Attribution: Perceiving the Causes of Behavior (Morristown, N.J.: General Learning Press, 1972)
- C. E. Osgood and P. H. Tannenbaum, "The Principle of Congruity in the Prediction of Attitude Change," Psychological Review, 62(1955), 42-55.
- M. Riess and B. R. Schlenker, "Attitude Change and Responsibility Avoidance as Modes of Dilemma Resolution in Forced Compliance Settings," Journal of Personality and Social Psychology, 35(1977), 21-30.
- T. S. Robertson, Consumer Behavior (Glenview, Ill: Scott, Foresman, 1970)
- M. J. Rosenberg, "Cognitive Structure and Attitudinal Affect," Journal of Abnormal and Social Psychology, 53 (1956), 367-372.
- B. R. Schlenker, "Self-Image Maintenance and Enhancement: Attitude Change Following Counterattitudinal Advocacy," Proceedings of the 81st Annual Convention of the American Psychological Association (Montreal, 1973)
- B. R. Schlenker, "Self-Presentation: Managing the Impression of Consistency When Reality Interferes With Self-Enhancement," Journal of Personality and Social Psychology, 32(1975), 1030-1037. (a)
- B. R. Schlenker, "Group Members' Attributions of Responsibility For Prior Group Performance," Representative Research in Social Psychology, 6(1975), 96-108. (b)
- B. R. Schlenker, "Liking for a Group Following an Initiation: Impression Management or Dissonance Reduction?" Sociometry, 38(1975), 99-118. (c)
- B. R. Schlenker and D. R. Forsyth, "Social Identity Theory: The Effects of Incentives on Attitude Change Following Counterattitudinal Behavior," Manuscript submitted for publication (Available from the authors, University of Florida, 1977)
- B. R. Schlenker and R. S. Miller, "Egocentrism in Groups: Self-Serving Biases or Logical Information Processing?" Journal of Personality and Social Psychology (in press, 1977)
- B. R. Schlenker and M. Riess, "Self-Presentations of Attitudes Following Commitment to Proattitudinal Behavior," Manuscript submitted for publication (Available from the authors, University of Florida, 1977)
- B. R. Schlenker and P. A. Schlenker, "Reactions Following Counterattitudinal Behavior Which Produces Positive Consequences," Journal of Personality and Social Psychology, 31(1975), 962-971.
- B. R. Schlenker, S. Soraci Jr. and B. McCarthy, "Self-Esteem and Group Performance as Determinants of Egocentric Perceptions in Cooperative Groups," Human Relations, 29(1976), 1163-1176.
- W. A. Scott, "Attitude Measurement," In G. Lindzey and E. Aronson (Eds.), Handbook of Social Psychology, 2 (Reading, Mass.: Addison-Wesley, 1968)
- M. Sherif and C. I. Hovland, Social Judgment: Assimilation and Contrast Effects in Communication and Attitude Change (New Haven, Ct.: Yale University Press, 1961)
- C. W. Sherif, M. Sherif and R. E. Nebergall, Attitude and Attitude Change: The Social Judgment-Involvement Approach (Philadelphia, Pa.: W. B. Saunders, 1965)
- M. L. Snyder, W. G. Stephan and D. Rosenfield, "Egotism and Attribution," Journal of Personality and Social

Psychology, 33(1976), 435-441.

I. D. Steiner, "Reactions to Adverse and Favorable Evaluations of Oneself," Journal of Personality, 36(1968), 553-564.

S. E. Taylor, "On Inferring One's Attitudes From One's Behavior: Some Delimiting Conditions," Journal of Personality and Social Psychology, 31(1975), 126-131.

J. T. Tedeschi, B. R. Schlenker and T. V. Bonoma, "Cognitive Dissonance: Private Ratiocination or Public Spectacle?" American Psychologist, 26(1971), 685-695.

W. I. Thomas and F. Znaniecki, The Polish Peasant in Europe and America, 1(Boston: Badger, 1918)

R. J. Wolosin, S. J. Sherman and A. Till, "Effects of Cooperation and Competition on Responsibility Attribution After Success and Failure," Journal of Experimental Social Psychology, 9(1973), 220-235.

C. B. Wortman, P. R. Costanzo and T. R. Witt, "Effects of Anticipated Performance on the Attributions of Causality to Self and Others," Journal of Personality and Social Psychology, 27(1973), 372-381.

R. S. Wyer Jr., Cognitive Organization and Change: An Information Processing Approach (Potomac, Md.: Lawrence Erlbaum Associates, 1974)

P. G. Zimbardo, The Cognitive Control of Motivation (Glenview, Ill.: Scott, Foresman, 1969)

A FUNCTIONAL APPROACH TO CONSUMER ATTITUDE RESEARCH

Richard J. Lutz, University of California, Los Angeles¹

Abstract

The potentials and current limitations of the so-called functional approach to the study of attitudes are delineated. In order to more fully illuminate the functional approach, the evolution of its research tradition is traced. The core of the functional approach, it is argued, can be captured by an expectancy-times-value reconceptualization of the theory. Implications for research and practice under the new conceptualization are discussed.

Introduction

It has been over fifteen years since Daniel Katz published his classic article, "The Functional Approach to the Study of Attitudes" (Katz, 1960). In this paper, which was essentially the last in a series of studies espousing the functional approach, Katz outlined the basic notion of the approach--i.e., that people hold attitudes toward objects, events, issues and behaviors for various reasons. That is, attitudes fulfill functions for the individual, such as maximizing rewards, expressing one's values, etc. He went on to specify four basic functions that an attitude may perform for an individual--utilitarian, value-expressive, ego-defensive and knowledge--and presented a summary of conditions under which the various functions might be aroused and another set of conditions conducive to the change of attitudes serving the four functions.

Katz's article was a major conceptual contribution to the attitude literature in that it made explicit the motivational underpinnings of attitudes and drew attention to the possible individual differences (in attitude function) which may be obscured by a simple affective representation of attitude. Consumer behavior scholars have been quite cognizant of Katz's contribution, as evidenced by the frequency with which the functional approach is discussed in consumer behavior texts. Virtually every major book in the field includes some treatment of the functional approach, and generally concludes that a functional approach is fundamental to understanding (1) why people hold the attitudes they do, and (2) how those attitudes might be changed.

Unfortunately, consumer researchers, and attitude researchers in general, have not shown as much interest in actually applying the functional approach. In fact, recently the functional approach has come under attack by some researchers who cite the lack of operational procedures to allow its application (Kiesler, Collins and Miller, 1969; Day, 1973; Eagly and Himmelfarb, 1974). The currently fashionable position on the functional approach, therefore, is that it is "...more important as a classificatory schema and general analytical framework than as a theory generating hypotheses for research" (Eagly and Himmelfarb, 1974. p. 598).

The purpose of the present paper is to develop a rationale for the measurement of attitude functions. To accomplish this, a review of the historical development of the functional approach will be undertaken, followed by a proposed reconceptualization. Implications of the reconceptualized functional framework for consumer attitude research will also be discussed.

¹Associate Professor of Marketing, UCLA. The present paper is an excerpt from a more comprehensive version (Lutz, 1977) which may be requested from the author.

Evolution of the Functional Approach

The functional approach to attitudes subsumes two independent schools of thought, one fostered by M. Brewster Smith and his colleagues at Harvard, and the other represented by Katz, Sarnoff and others at the University of Michigan. Though developed independently, the two approaches are remarkably similar. Each will be reviewed briefly here, in order to highlight their similarities and differences, and to provide a flavor of the functional approach.

The Harvard Study

At the close of World War II, the American people were confronted with the emergence of Russia as a major world power. Communism was a dreaded menace, and the concern among the public was great. Smith and his colleagues set out to study the nature of attitudes toward Russia and, in particular, the reasons underlying these attitudes.

Smith (1947), in the first explicit statement of functional theory, posited a relationship between attitudes and "personality." Based on the results of a survey of 250 New England males, Smith identified five functions which attitudes could serve for the personality:

- 1) the value function, wherein the attitude expresses a value of the individual;
- 2) the consistency function, which maintains the relationship between attitudes and characteristic modes of reaction (e.g., personality traits);
- 3) the gratification function, whereby the attitude indirectly satisfies some basic need;
- 4) the meaning function, which allows the person to give order and stability to his world; and
- 5) the conformity function, which is served by attitudes expressing identity with, and fostering acceptance by, important social groups.

The essential thrust of Smith's functional formulation, then, rested on the postulate that attitudinal affect is in some way based on the relationship of the attitude object to underlying psychological constructs variously labeled values, motives and needs. The partitioning into five distinct functions is perhaps best seen as an attempt to specify some qualitative differences among the underlying needs. For example, the meaning function was seen as relating to the needs for "adequacy" and "simplicity," while the conformity function centered on "need for approval" and "need for acceptance." The value function dealt with "what sorts of things people think are important in life" (Smith, 1947, p. 518) such as liberty and economic security, while the consistency and gratification functions analyzed the relationship between attitude and traditional personality variables (e.g., extroversion).

While Smith (1947) did not go into any detail regarding the mechanisms through which the various functions operate, it seems clear that the indirect satisfaction of needs as specified by the gratification function would differ qualitatively from more direct mechanisms. Thus, there would appear to be not only differences in the content of needs across the five functions, but also qualitative differences in the psychological mechanisms

implied by the functions.

In a subsequent paper, Smith amplified on the nature of the value function:

A person will tend to perceive and judge the focus of an attitude in terms of one of his personal values to the extent that (a) the value is important to him, occupying a central position in his value hierarchy; (b) the information available to him about the focus contains a basis for engaging the value; and (c) the scope of the value and of the person's interest is broad enough to extend to the focus of the attitude. (Smith, 1949, p. 486).

Thus, although the other functions were not similarly analyzed, Smith did engage in theorizing with respect to the conditions necessary (and presumably sufficient) for the value function to be an underlying mechanism in support of an attitude. This represents at least a partial rebuttal to those who contend that functional theory is not really "theory," but rather is little more than simply a taxonomy of attitude-related needs and values (Eagly and Himmelfarb, 1974; Baron, 1968).

In 1956 Smith, Bruner and White authored Opinions and Personality, which reported on the entire study of attitudes toward Russia, but focused primarily on the results of in-depth interviews with 10 men. From this study emerged a modification of Smith's earlier functional theorizing, together with an elaboration of theoretical mechanisms embodied by the functional approach.

Rather than the five functions postulated by Smith (1947), Smith *et al.* (1956), conceived of three basic functions: object appraisal, social adjustment and externalization. The latter function is closely related to Smith's earlier gratification function, while social adjustment corresponds to his conformity function. The object appraisal function has no direct analogue in Smith's earlier work, although it does appear to include his meaning function. Oddly, the value function, which was so central in Smith's earlier work, seemingly is ignored in the Smith, *et al.* (1956) statement.

The basis for the object appraisal function is "reality testing," in that an attitude serving the function is seen as classifying the attitude object for the purposes of allowing the individual to adequately cope with the object. Stated differently, "...object appraisal is the process whereby the person develops attitudes that are a creative solution to the problems posed by the existence of disparate internal demands and external or environmental demands" (Smith, *et al.*, 1956, p. 41). Again emphasizing the basic need-relatedness of attitudes, Smith, *et al.* further specify that objects are appraised in terms of their relations to the person's "motives, goals, values and interests" (p. 261). But they are careful to point out that they are not advocating "...that admirable fiction, the 'rational' man," (p. 265), but rather are postulating object appraisal as only one of the functions attitudes may serve.

Smith, *et al.*'s social adjustment function deals with "...facilitating, disrupting or simply maintaining an individual's relations with other individuals (p. 41). Such processes as identification, conformity, which represents a broadening of Smith's (1947) conformity function.

Finally, the externalization function is operating when "...an individual, often responding unconsciously, senses an analogy between a perceived environmental event and some unresolved inner problem" (Smith, *et al.*, 1956, p. 43). This externalization helps the individual

to reduce anxiety. Like the other functions the degree and importance of externalization are expected to differ greatly across both individuals and situations.

Turning to an analysis of attitude change implications of the functional approach, Smith *et al.* state that while any one function can be a "precipitating factor" in attitude change, the process of change must entail a "...shift in the balance of all factors" (p. 276). This implicitly assumes a weighting of the various functions in their relationship to attitude, and this weighting has important implications for the form that attitude change strategies might take. For instance, a predominantly object appraisal attitude would most likely be modified by the presentation of information, while an attitude serving the social adjustment function ought to be more susceptible to "prestige suggestion" (Smith, *et al.*, 1956, p. 277). Attitudes which are primarily based on externalization are expected to be more rigid and not subject to modification through either information or social manipulation (p. 277). As Smith, *et al.* conclude:

If this analysis of the conditions of attitude change is correct...it becomes important to devise rigorous and objective means of assessing the contribution of the different functional supports to people's opinions on given issues. Ways of doing so seem entirely within the range of feasible attainment." (p. 278)

But Smith, Bruner and White did not attempt to develop the measures they specify, leaving these important attitude change propositions untested and largely untestable. This lack of operational procedures remains the single most glaring weakness of functional theory, not only for Smith, *et al.*, but also for the work of Katz and his associates at Michigan.

The Michigan Group

Just as the Harvard Study had centered on an important social phenomenon--attitudes toward Russia--the University of Michigan group focused its work on the problem of understanding and changing prejudicial attitudes. The earlier work of Adorno, *et al.* (1950) on the authoritarian personality and its relationship to anti-Semitism had a significant impact on the Michigan group. However, rather than concentrate solely on the personality antecedents of prejudice, the Michigan group broadened its study to include a number of motivational bases supportive of attitude formation and change.

In the first paper in the series Sarnoff and Katz (1954) outlined three "motivational contexts" for attitudes corresponding roughly to the major psychological research traditions of Gestalt theory, learning theory, and psychoanalytic theory. These three contexts were labeled reality testing, reward and punishment, and ego-defense, respectively.

Reality testing was described in terms of the need to explore, the need to know, the quest for information, curiosity and the like. In what today might be described as an information processing view, Sarnoff and Katz (1954) stated that attitudes serving this function are "...a function of the range of information which has been accessible to the individual in regard to certain target objects" (p. 117). Further, they specified that such attitudes could be changed via a "rational" approach involving the presentation of new information within an appropriate frame of reference.

Attitudes based on reward and punishment were seen as an amalgam of two basic sorts of rewards: (1) those based on social group norms, and (2) those emanating from the individual's internal value system. Sarnoff and Katz were more interested in the former type (because of the

presumed social nature of prejudice) and therefore emphasized the use of normative group pressure in changing such attitudes. In addition, however, they noted that restructuring the value system "...through making the specific connections between the target object and a different set of values" (p. 122) was also a viable modification strategy.

Finally, attitudes based on ego-defense were viewed as serving "consciously unacceptable" motives through such ego-defense mechanisms as projection. Ego-defensive attitudes were regarded by Sarnoff and Katz as "inappropriate," in that the cognitive object...perceived contrasts with characteristics of the object" (p. 120). These distortions could only be corrected through attitude change procedures involving some form of catharsis, which would allow the individual to perceive and respond to reality in a more "appropriate" manner.

Similar to Smith, et al. (1956), Sarnoff and Katz (1954) noted that any given attitude may serve one or a combination of functions. They concluded their paper with a discussion of attitude change "programs" designed to modify attitudes serving the various functions. While they reported no empirical work, a number of studies followed in the next several years which were aimed at investigating some of Sarnoff and Katz' ideas.

Reality testing. Cohen, Stotland and Wolfe (1955) investigated the need for cognition (n-cog) as "...the individual's tendency to organize his experience meaningfully" (p. 291). This corresponds directly to Sarnoff and Katz' (1954) reality testing function. Using two independent measures of n-cog, they were able to demonstrate that students' feelings of "frustration" were greater when reading "ambiguous" stories than when reading unambiguous stories, and further, that the increment in frustration was greater for students who were high in n-cog. While they failed to support two hypotheses regarding the degree to which the students would attempt to "impose structure" on the ambiguous stories, their study was important in documenting that n-cog did exist and that affective reactions to a situation varied with both intrapersonal (i.e., n-cog) and environmental (i.e., manipulated ambiguity) factors.

In a subsequent study, Cohen (1957) investigated the relationship of n-cog to the order of presentation in persuasion. Specifically, he used two orders of presentation, one in which need-arousal was followed by information, with another which reversed the two. His hypothesis was that those subjects high in n-cog would be less affected by the order of presentation, since they would work harder to structure the situation for themselves. Using a forced-choice scale to measure n-cog, he divided a class of 35 students into high and low groups. A communication supporting "grading on the curve" was presented either before or after a "fear-arousing" discussion of the vagaries of the grading system.

Cohen (1957) found greater attitude change on an immediate posttest for the group receiving the need-arousal treatment first. He concluded:

"Need satisfaction and presumably, therefore, opinion change in this situation demanded the perception of the information as instrumental to the aroused need" (p. 94).

However, on a delayed posttest, Cohen found that the high n-cog subjects retained a significant change in attitude regardless of which group they were assigned to, while the low n-cog subjects showed no positive change, regardless of which order of communication they had received. The results of this study served to re-affirm the presence of a reality testing function and its mediating influence on attitude change.

Reward/punishment. In 1955, Helen Peak published her classic article, "Attitude and Motivation," in which she outlined an expectancy-times-value (ExV) approach to representing the motivational structure underlying attitudes. Peak was a member of the Michigan group and a part of the Attitude Change Project under Katz' direction, and she assumed directorship of the project for one year while Katz was away. During this time, a number of studies were conducted which, while closely related to the functional approach, did not explicitly position themselves within the Sarnoff-Katz functional framework.

In the first of these studies, Rosenberg (1956) investigated the "attitude structure" underlying attitudes toward free speech for Communists and the removal of Negro housing segregation. By measuring the "value importance" of each of 35 goals and values, as well as the "perceived instrumentality" of the two issues mentioned above in blocking or attaining those goals, Rosenberg was able to construct an ExV "cognitive structure index" which was closely related to overall affect. Rosenberg's work was later cited by Katz, et al. (1956) as "...a combination of the rational and reward approaches..." (p. 28). The "rational" portion derived from the measure of instrumentality, while the "reward" function was represented by the value importance hierarchy.

In a second study, Carlson (1956) built on Rosenberg's earlier work on Negro housing segregation in attempting to modify instrumentalities relating to the values of American prestige, protection of property values, equal opportunity, and being broadminded and worldly-wise. By having students write essays supporting the removal of segregation and then following this with an in-class discussion, Carlson was able to modify the individual instrumentalities, the overall ExV index, and to a somewhat lesser extent, attitude. Subjects with initially moderate attitudes changed the most while initially favorable subjects had little room to change, and initially negative subjects exhibited a "boomerang" effect. This latter finding suggests that some of Carlson's subjects were holding ego-defensive, rather than reward/punishment attitudes, but since Carlson included no measure of ego defense, the finding is only suggestive. Katz, et al. (1956) also cite Carlson's (1956) study as utilizing a combination of the reward and rational approaches, although it should be noted that neither Rosenberg (1956) nor Carlson (1956) characterized their own work in this way.

In the subsequent studies reported by Peak and her associates (Peak and Morrison, 1958; Peak, 1960; Rosenberg, 1960), the emphasis continued to be placed on the ExV representation of cognitive structure, with no explicit links being made to the functional approach.

Ego-defense. Wagman (1955) tested Sarnoff and Katz' (1954) propositions regarding the ego-defense function by attempting to modify students' anti-Negro attitudes through both authoritarian suggestion and a "non-authoritarian" cognitive restructuring technique. Splitting his sample into thirds based on their F-Scale scores, Wagman was able to show a fairly consistent pattern across four different attitude measures:

- 1) low F-Scale subjects changed more than high F-Scale subjects under conditions of cognitive restructuring;
- 2) high F-Scale subjects changed more than low F-Scale subjects under conditions of authoritarian suggestion, regardless of which direction of change was advocated (i.e., more anti-Negro prejudice or less anti-Negro prejudice).

These results are consistent with the Sarnoff and Katz' (1954) reasoning that ego-defensive (in this case, authoritarian) attitudes will be less susceptible to

change through a "rational" approach. However, Wagman's findings indicate that ego-defensive attitudes are not necessarily rigid in the face of less "rational" attitude modification strategies.

In two related studies which were aimed at directly investigating the ego-defensive function of attitude, Katz, Sarnoff and McClintock (1956) and Katz, McClintock and Sarnoff (1957) attempted to change attitudes toward Negroes. In the first study, both an informational (rational) and "interpretive" (ego-defensive) appeal were used, in order to determine their relative effectiveness in attacking a prejudicial and presumably ego-defensive attitude. In order to measure ego-defense, a TAT card was designed and codes were developed for projection, denial and extrapunitive, three ego-defense mechanisms. These were combined into a single score which was used to split the sample into thirds. It was predicted that the efficacy of the informational appeal would decrease with increasing ego-defense, while the interpretive appeal would be maximally effective for moderately defensive subjects. Neither prediction was confirmed, which led Katz *et al.* to question the ego-defense measure they had employed.

In their second study, Katz, *et al.* (1957) used four separate measures of ego-defense: a pair of TAT cards, an 18-item multiple choice sentence completion task, 45 items from the paranoia subscale of the MMPI, and six subscales from the F-Scale. The former two measures were found to be insensitive, while the latter two were judged to perform adequately. An interpretive appeal was administered to 131 college women, with the expectation that those moderate in defensiveness would exhibit more change than either the high or low defensiveness groups. This prediction was confirmed for the two standardized measures (MMPI and F-Scale), while being disconfirmed for the two more *ad hoc* approaches.

McClintock (1958) investigated the relationship between attitude change and two personality variables, ego-defensiveness and other-directedness. One hundred ninety-eight female college students were exposed to one of three forms of persuasive appeal (information, insight, ethnocentric) designed to change anti-Negro attitudes. Consistent with Katz, *et al.* (1957) McClintock found that moderately defensive subjects (as measured by a portion of the F-Scale) changed the most under the insight treatment, while no differences were observed among the groups under the informational appeal. Under the ethnocentric appeal, subjects high in ego-defense changed more than those low in defensiveness. McClintock (1958) concluded:

...the findings of this research indicated that a given influence may have little effect in moving the total population in the direction of that influence... Thus, knowing the personality structure of the individual seems necessary for... the devising of appropriate and effective methods for changing attitudes. (pp. 492-3).

In yet another study of the self-insight approach to changing ego-defensive attitudes, Stotland, Katz and Patchen (1959) attempted to manipulate racial attitudes. Contrary to the pattern of findings emerging from earlier studies, a monotonic decreasing relationship was obtained between ego-defense (as measured by the F-Scale) and attitude change. This failure to replicate the earlier findings was not readily explainable, but Stotland, *et al.* speculated that it could be due to an insensitive measure of ego-defense (even though the same measure had been used satisfactorily in a number of other studies), increased awareness by low defensive subjects of their "inappropriate" attitudes, or a greater amount of projection among those low in ego-defense than had been previously anticipated. This study serves to underscore the rather weak operational procedures characterizing the

research on the ego-defensive function, even though this function was clearly the major focus of empirical research by Katz and his associates.

In his summarization of the ego-defensive function of attitude Sarnoff (1960a) states:

It begins to appear that--insofar as consciously unacceptable motives do form the basis of an individual's disposition to agree with prevailing anti-minority stereotypes--projection plays a larger role than any of the other mechanisms of ego defense. (p. 273)

He concluded this on the basis of the results of a number of studies utilizing projection and other defense mechanisms (Katz, *et al.*, 1956, 1957; Sarnoff, 1951, 1960b; Sarnoff and Corwin, 1959; Cohen, 1956). Thus it would appear that any further attempt to investigate the ego-defense function of attitude should rely heavily on projection as a basis for assessing attitudinal function.

In the first major theoretical statement of functional theory since the Sarnoff and Katz (1954) article, Katz and Stotland (1959) outlined a theory of attitude formation and change based upon the research described above. At the heart of their theory were the motivational functions served by attitudes. These functions were classified as proximal, object-instrumental, ego-instrumental, and ego-defensive.

Proximal attitudes are those where the attitude toward the object satisfies some need of the individual directly. This is in contrast to attitudes which are held toward an object because it is instrumental in helping the individual to satisfy a need. Attitudes toward foods and sex objects are cited as possible examples of proximal attitudes, but perhaps the most important example is the attitude based on the need for understanding. This is directly analogous to Cohen's (1955, 1957) investigations of n-cog and re-asserts the notion that to understand, to explore, to search for meaning is an end in itself, and not merely a means to some other end-state. Proximal attitudes are seen as being reinforced through repeated interaction with the attitude object and are "...based on the principle that individuals put high value on objects which satisfy their needs and low value on objects which frustrate them" (Katz and Stotland, 1959, p. 437).

Object-instrumental attitudes are those based on the means-ends principle first explicitly stated by Peak (1955) and her associates. According to Katz and Stotland (1959), many social attitudes fall into this category. Positive attitudes are developed toward those social objects which are seen as a means to ultimate satisfaction of wants and needs. These sorts of attitudes may be learned through trial-and-error, but as noted by Katz and Stotland: "The concept of instrumental learning has been converted into the notion of perceived instrumentality by workers dealing with cognition, perception and attitudes" (p. 436). Thus, the characteristics of the attitude object and their relationship to need satisfaction act as the motivational underpinnings of object-instrumental attitudes.

In contrast, an ego-instrumental attitude "...arises from sources further removed from the attitude itself, from ego satisfactions" (p. 440). The key feature of this type of attitude is that it allows the person to express to others what kind of person he is. Thus, such constructs as "value" and "self concept" become relevant attitude bases. When compared with object-instrumental attitudes, the cognitions underlying ego-instrumental attitudes are "...elaborated less upon the basis of the objective characteristics of the attitudinal object than upon the individual's need to maintain his own self-image" (Katz and Stotland, 1959, p. 440).

The final category of attitudes, ego-defensive, is essentially the same as was discussed by Sarnoff and Katz (1954). Ego-defensive attitudes conform to the classic pattern of repressing consciously unacceptable sexual and aggressive impulses and subsequently projecting these impulses onto others, usually outgroups of some kind. In this sense, ego-defensive attitudes are maladaptive and leave the basic conflict within the individual unresolved. They are quite resistant to change through common forms of influence, which follows from the deep-seated nature of the motives being repressed.

Katz and Stotland viewed the motivational-attitudinal-action system as dynamic, characterized by constant interaction, such that no true distinctions can be made among independent and dependent variables. However, they did feel that certain variables would be more dominant in a causal sense:

...we would in general regard motives and environmental forces as independent variables, attitudes as intervening variables, and expression in behavior as the dependent variable (p. 468).

This view of the causal structure of the system has direct implications for the derivation of attitude change strategies; hence, Katz and Stotland recommend modifications of (1) the individual's value system; or (2) cognitions with respect to the attitude object as the most viable means of changing attitudes.

In the final major paper pertaining to functional theory, Katz (1960) offered what is perhaps the most systematic description of the proposed functions as well as the conditions conducive to their arousal and change. The four functions he outlined were similar, but not identical, to the ones described by Katz and Stotland (1959): utilitarian (also referred to as instrumental or adjustive), ego-defensive, value-expressive, and knowledge.

The utilitarian function was described by Katz as an attempt by the individual to maximize rewards and minimize punishment from the environment. Thus, attitude formation was seen as "...dependent upon present or past perceptions of the utility of the attitudinal object for the individual" (p. 171). This relates closely to the notion of object-instrumental attitudes presented by Katz and Stotland (1959).

The ego-defensive function, as conceived by Katz (1960), is essentially identical to the one originally conceived by Sarnoff and Katz (1954) and reiterated by Katz and Stotland (1959). It is through this function that the individual protects himself from either external threats or internal conflicts. In its linkage to Freudian psychology, the ego-defensive function is concerned primarily with consciously unacceptable motives.

The value-expressive function involves attitudes which express the individual's own personally-held values or his self concept. This is analogous to Katz and Stotland's (1959) ego-instrumental function. It is virtually opposite to the ego-defensive function, which serves to obscure aspects of the person from himself and others. Value-expressive attitudes instead reveal the self concept and in some cases help the individual to move closer to his "ideal" self concept.

The knowledge function is served by attitudes which help the individual to find meaning, organize the environment, and in general provide clarity and consistency to the person's view of the world. This is in agreement with the proximal attitude based on n-cog as described by Katz and Stotland.

Katz' (1960) major contribution was a formal specification of the factors underlying attitude formation, and change. His was the only systematic analysis of the nature of the various functions, together with an assessment of the key elements which could be brought to bear on attitude change efforts. While some later authors have criticized the conditions he outlined as being too general to allow empirical test, nevertheless the formal statement of these conditions was at least a good first step toward building a set of testable propositions, which is a requirement for any theory.

More disappointing than the generality of the determinants of attitude arousal and change was the lack of any clues as to how the approach might be operationalized. Katz (1960) cited the need for public opinion polls measuring needs and attitudes, but beyond a few references to the F-Scale, MMPI and n-cog, he gave no indication as to how the functions could be assessed. It would appear that at least some of the difficulty that arises in the assessment of attitude functions derives from a lack of conceptual clarity with respect to the various functions. By moving toward appropriate measurement procedures, some of the somewhat vague notions regarding attitude arousal and change conditions may crystallize into more clearly testable hypotheses.

Summary

Table 1 summarizes the various functions of attitude discussed in five major theoretical papers. As is evident from the table, there are some discrepancies among the articles in terms of both the number of functions and the exact nature of these functions. The two functions which were identified with the most consistency across the various studies were (in Katz' (1960) terminology) the ego-defensive and knowledge functions, although the latter was characterized by five different labels and in one paper (Smith, et al., 1956) was combined with the equivalent to Katz' (1960) utilitarian function into a single "object appraisal" function. The consensus on the ego-defensive function is not surprising, given the involving and apparently threatening nature of the attitudes under study (i.e., attitudes toward Russia and racial and ethnic minorities).

With the exception of Smith's (1947) article, there was agreement on a utilitarian function stressing the reward/punishment notion of hedonism. Somewhat less agreement was found, however, on the existence of a value-expressive function. It appeared in the later work of the Michigan group, but was not considered by either Sarnoff and Katz (1954) or by Smith, et al. (1956). In fact, the latter authors explicitly rejected the notion of a value-expressive function (although Smith (1947) had earlier postulated one), and instead argued that value expression was not a "need" of the individual but rather just a by-product of the expression of an attitude (Smith, et al., 1956, p. 38).

The clearest difference between the Harvard and Michigan groups is apparent in their treatment of the social adjustment function. Smith, et al. (1956) strongly emphasized the social function of attitude, while Katz and his associates (1954, 1959, 1960) did not distinguish it as a function per se. To some degree, the social function of attitude was subsumed in Katz' (1959, 1960) discussion of the value-expressive function, but the discussion focused more on socialization as a source of values than on social needs as direct underpinnings of attitude.

While there was some disagreement as to the exact functions served by attitude, the theorists were unanimous in their characterization of attitudes as having multiple determinants. Although it was conceded that unifunctional attitudes may exist in some cases, the consensus was that most attitudes serve more than one function

TABLE 1

A COMPARISON OF THE FUNCTIONS OUTLINED IN MAJOR THEORETICAL PAPERS

Basic Need Syndrome	Smith (1947)	Smith, et al. (1956)	Sarnoff & Katz (1954)	Katz and Stotland (1959)	Katz (1960)
Expression of values or self concept	value	a	_____	ego-instrumental	value-expressive
Indirect gratification of consciously unacceptable motives	consistency gratification	externalization	ego-defense	ego-defensive ^c	ego-defensive
Satisfaction of social needs	conformity	social adjustment	_____	_____	_____
Search for knowledge, organization consistency	meaning		reality testing	proximal ^d	knowledge
Maximization of reward or minimization of punishment	_____	object appraisal ^e	reward and punishment	object-instrumental	utilitarian

^a not explicitly a function of attitude, but rather a by-product

^b can also relate to "consciously acceptable" personality traits

^c also can be regarded as exhibiting properties of a proximal attitude.

^d is not limited to the need for understanding

^e combined the ideas of reward/punishment and need for understanding

for the individual, and that only through an understanding of these functional bases could attitudes be diagnosed and effectively changed. Thus, the functional theorists agreed on an approach to attitude which was strongly based on individual differences; i.e., the same observed attitude (in terms of strength and direction of affect) for two different individuals may be reflecting two entirely different motivational structures in terms of the functions being served.

A Reconceptualization

The lack of operations for measuring functions of attitude has been identified as a major drawback to functional theory (e.g., Insko, 1967; Himmelfarb and Eagly, 1974). Some researchers have gone even further with this criticism suggesting that the theory is really nothing more than a broad conceptual integration and should not be viewed as a potential operational framework (Kiesler, Collins and Miller, 1969). However, the functional theorists themselves clearly regarded the theory as potentially operationalizable. Smith, et al. (1956), while offering no clues as to how to operationalize the theory, nevertheless pointed out the need to do so and voiced optimism that the measurement of attitude functions "...seem entirely within the range of feasible attainment" (p. 278). Similarly, Katz (1960) discussed the assessment of attitude functions, pointing to the need for systematic sampling of target populations and advocating increasing reliance on methods other than the depth interview.

Yet, despite the early attempts to measure ego-defensive (e.g., Katz, et al., 1956, 1957) and knowledge function

(e.g., Cohen, et al., (1955) attitudes, little systematic empirical work was devoted to the task of functional assessment. Perhaps the best explanation for this problem lies in the following quotation from Katz and Stotland (1959):

...we cannot systematically develop the conceptual properties of all our constructs, point to validated operational measures for them, or describe an appropriate mathematical model for handling the data in this field. (p. 425, italics added)

Thus, although the Katz and Stotland (1959) contribution was considered only a "preliminary" statement of functional theory, the final conceptual article was published the very next year by Katz (1960). In neither of these articles was there sufficient conceptual clarity to allow rigorous specification of operations for representing the various functions and their relationship to the expressed attitude of an individual. Yet, the kernel of a fundamental approach to solving the conceptual, and, concomitantly, the operational problems with functional theory was present in these and earlier articles. The following section attempts to explicitly draw out from the functional literature the roots of the proposed reconceptualization.

An Expectancy-Times-Value View²

Weiner (1972) describes the expectancy-times-value (ExV) approach to human motivation as the conception that "...direction and intensity of behavior is a function of the expectation [E] that certain actions will lead to the goal, and the incentive value [V] of the goal object" (p. 8). The ExV notion is central to Murray's (1938) theory of personality, Rotter's (1954) social learning theory, Atkinson's (1964) theory of achievement motivation, and Peak's (1955) theory of attitude. A careful examination of the literature on the functional theory of attitude suggests that it, too, is strongly grounded in ExV principles.

In their discussion of the relationship of attitude to personality, Smith, Bruner and White (1956) posited that personality is marked by an "...adaptive striving after goals [V]," and further that the individual becomes "...selectively aware of objects related [E] to certain goals..." (p. 32). In a similar vein, Katz (1960) states:

Knowledge of need state [V] indicates the type of goal toward which the individual is striving. But the means for reaching this goal [E] may vary considerably...(p. 203).

Thus, the general notion of goal-oriented behavior is apparent in both of the major approaches to functional representation of attitude.

The basic idea of an ExV reconceptualization of functional theory is that the cognitive "component" of attitude serving a particular function becomes elaborated around the basic needs or values salient to that attitude function. The "index of cognitive structure" (ExV) becomes an index of the particular attitude function under consideration. The exact nature of the cognitive (E) and affective (V) elements making up the index would vary with the content of the various functions. A consideration of the individual functions serves to further illuminate this ExV orientation of functional theory.

Utilitarian.

Probably the clearest expression of functional theory's implicit reliance on ExV principles is found in Katz' (1960) discussion of the utilitarian or adjustment function:

Attitudes acquired in the service of the adjustment function are either the means for reaching the desired goal or avoiding the undesirable one....In general, then, the dynamics of attitude formation with respect to the adjustment function are dependent upon present or past perceptions [E] of the utility [V] of the attitudinal object for the individual...The closer these objects are to actual need satisfaction [V] and the more they

are clearly perceived [E] as relevant to need satisfaction, the greater are the probabilities of positive attitude formation (p. 171, italics added).

According to the ExV formulation, then, the utilitarian function would be measured by expectations regarding the need-related attributes of the attitude object, weighted by the value or utility of these attributes to the individual. The degree to which the utilitarian ExV index relates to attitudinal affect is an assessment of the degree to which the attitude is serving the utilitarian function for the individual.

Value-expressive. As noted by Katz and Stotland (1959), the utilitarian function rests primarily on the properties of the attitude object and their relationship to goal attainment. On the other hand, value-expressive attitudes are seen as reflections of properties of the individual:

The cognitive component of the attitude becomes elaborated less upon the basis of the objective characteristics of the attitudinal object than upon the individual's need to maintain his own self-image (Katz and Stotland, 1959, p. 440).

Katz (1960) postulates that the expression of the self concept is rewarding to the individual. Presumably, in order for an attitude to serve the value-expressive function, then, there must be some expectation (E) on the part of the individual that holding that particular attitude gives expression to some aspect of self (V). Or, as stated by Smith (1947):

...intensity of attitudinal affect is a function of the extent to which a personal value is engaged [E] and of the importance [V] of this engaged value in the hierarchy of the person's central values (p. 519).

Whether dealing with aspects of the self concept or central values of the individual, the mechanism underlying the value-expressive function relies on the perceived relationship (E) of the attitude object to some characteristic or value, and the strength or importance (V) of that value to the individual. The strength of the relationship between an ExV index based on individual values and overall attitude serves as an indicator of the value-expressiveness of the attitude.

Knowledge. In an investigation of the knowledge function, Cohen, et al. (1955) implicitly used an ExV conceptualization in their hypotheses that:

1. An ambiguous situation [E] will be more frustrating [i.e., negative affect-producing] than a more structured one.
2. Hypothesis 1 will be more applicable to people with high need for cognition [V] than for those with low need. (p. 291)

Need for cognition was postulated as a desired end-state (V) which could be satisfied by perceptions (E) of the environment as unambiguous and structured. The assumption of multiplicative combination of the E and V terms is clearly present in Hypothesis 2. In their later discussion of the knowledge function, Katz and Stotland (1959) state:

Thus it would be expected that those objects in the environment which aid in understanding the world would be evaluated highly. Furthermore, if the object itself is clearly understood, it will be evaluated more highly than if it is understood but vaguely (p. 438).

²In the recent discussion, "expectancy" will be used to refer to any perceived relationship between an attitude object and some related attribute, property or value. No distinction is made between the level of the perceived relationship (e.g., degree of sweetness of a soft drink) and the probability that the level exists (e.g., it is very likely that the soft drink is sweet). While this distinction may be important in subsequent measurement of attitude functions, it is not critical to the conceptual development of an ExV approach.

Therefore, it is apparent that the perception of the environment as unambiguous, or as aiding in the reduction of ambiguity, is the major factor in satisfying the need for knowledge function attitudes. An ExV index based on needs for knowledge, understanding, cognition, etc., and the related environmental perceptions would form a measure of the knowledge function. The relationship between this index and attitude would indicate the degree to which the knowledge function was being served by the attitude.

Ego-defensive. The final function to be considered here, ego-defensive, is on the surface rather far-removed from an ExV formulation in its emphasis on unconscious motives. In fact, Locke (1975) criticized ExV theory as being incompatible with subconscious motives. However, there is reason to believe that this criticism may be somewhat overstated, in view of several observations made by functional theorists. For instance, in discussing their externalization function, Smith, et al. (1956) postulated that:

Covert strivings influence selectivity in the perception [E] of objects [p. 271] ...Attitudes toward objects are influenced by attitudes toward covert strivings [V] (p. 273).

Similarly, Sarnoff and Katz (1954), in discussing the distortion that is inherent in ego-defensive attitudes, noted that four types of distortion could be present:

In terms of distortion of the target object, the person with ego-defensive attitudes may:

1. Attribute to others [E] qualities which they do not possess.
2. Ignore [E] qualities which they do possess.

In terms of affect loading, he may:

1. Experience an excessive emotional reaction (either positive or negative) [V] to real or attributed qualities of the object.
2. Fail to experience an adequate emotional reaction (either positive or negative) [V] to real or attributed qualities of the target. (p. 120)

Finally, Sarnoff (1966) notes:

...a precise determination of the functional relationship between an attitude and a consciously unacceptable motive involves:

1. A postulation of which combination of consciously unacceptable motive [V] and ego defense [E] might plausibly account for the particular overt response from which the attitude is inferred.
2. After conceptualizing the most plausible combination of ego defense and consciously unacceptable motive, we must proceed to demonstrate empirically the relationship between that combination [ExV] and the attitude it is presumed to support. (p. 283).

It has been previously noted that the ego defense mechanism of projection is the only one which has been shown to consistently relate to social attitudes (Sarnoff, 1960a). To the extent that the projection of unacceptable traits and motives onto outgroup members constitutes an expectation (E), however inappropriate, then the resultant satisfaction of ego-defensive motives (V) should lead to

the formation of ego-defensive attitudes. Further, an ExV index composed of all projected traits and their associated evaluations should provide an index of the ego-defensive function. To the extent that this index is associated with a measure of overall attitude, the attitude can be regarded as serving the ego-defensive function.

The preceding references to, and quotations from, functional theory literature should point out the considerable compatibility of functional theory with an ExV formulation, although this linkage was not explicitly recognized by the functional theorists, nor by later attitude researchers. While it may be argued that the ExV notion is obvious in the writings of functional theorists, it is not so obvious as to have been carried through in systematic fashion. This is particularly evident in Katz' (1960) hypothesized conditions mediating attitude arousal and change and in the lack of operations for assessing attitude functions. It is therefore asserted that viewing functional theory from an ExV perspective constitutes a first step in reconceptualizing the theory along lines which explicitly note its compatibility with ExV theory and move it further toward the status of an empirically testable theory.

Implications

There are numerous important implications stemming from the proposed reconceptualization of functional theory. These implications can be broken down into implications for theory and implications for policy.

Theoretical Implications

Most importantly, the ExV reconceptualization of functional theory provides a means for empirical testing of functional theory propositions. Thus this conceptually rich approach to the study of attitude and attitude change can, for the first time, be thoroughly investigated to determine its usefulness as a theory.

Secondly, the ExV framework proposed here, together with appropriate operational procedures (Lutz, 1977), can perform an integrative function in at least two respects. At one level, the reconceptualization has helped to expose the basic similarity of process underlying the various functions. At another level, it can help to achieve the integration of multiattribute models of attitude (Wilkie and Pessemier, 1973), as well as self-concept research (e.g., Ross, 1971) and the Rokeach (1968) value paradigm, under the functional umbrella, an important unifying step in current research on attitudes. Self-concept research, in particular, has been a stepchild in consumer behavior research, with no broader frame of reference. Functional theory provides that frame of reference.

Policy Implications

Functional theory is a potentially powerful policy tool; until now, however, its usefulness was hampered by the lack of measurement procedures. The current approach should lead to a solution of the measurement problem, thus making the theory available for market application. Its contributions should be felt in at least three areas.

Attitude measurement. Functional theory, as reconceptualized, would provide a unified framework for the measurement of attitudes toward brands, products, social issues, etc. Depending upon the level of involvement, different weights should appear on the various functions. It should be of interest to the manager, for example, to know the functional basis of attitudes toward a particular brand in comparison with the functional basis of attitudes toward the product class in general. This

knowledge could lead to an industry-based campaign directed at one function, while brand advertising may focus on another function altogether.

Market segmentation. McClintock (1958) noted that "... knowing the personality structure of the individual seems necessary for the devising of appropriate and effective means for changing attitudes" (p. 493). Since modern segmentation theory espouses a "building up" of individuals into homogeneous market segments, functional theory is well-suited for market segmentation based on the particular attitude function(s) being served for various groups in the market. This approach to segmentation has been attempted, with some success, by Locander and Spivey (1975).

Communications strategy. Clearly, the functional approach should be quite useful in the diagnosis of consumer attitudes for the purposes of prescribing attitude change strategy. By measuring the various functional supports and determining their relative weights, the most efficient means of change can be derived. The functional framework also provides a basis for communications copy pretesting, to determine if the appropriate cognitive-motivational structure is being impacted. Finally, the proposed functional approach could be used for monitoring the effects of a persuasive campaign in the marketplace (Lutz, 1978).

Conclusion

In this paper, the origins and propositions of the functional theory of attitude have been reviewed, evaluated and reconceptualized within an expectancy-times-value framework. The proposed reconceptualization has been shown to be consistent with theoretical statements by the original functional theorists.

Functional theory is a potentially powerful managerial tool for use in advertising and segmentation decisions. Thus, adoption of the proposed ExV framework would not only enable testing of the theory, but also application of the theory to policy decisions. The potential richness of the theory for decision-making has long been noted but never attained due to lack of measurement procedures. The proposed reconceptualization does not completely ameliorate that deficiency in view of the fact that no data have been reported, but it does serve to make it clear that measurement is feasible within the reconceptualized functional framework.

References

T.W. Adorno, Else Frenkel-Brunswik, Daniel J. Levinson, and R. Nevitt Sanford, The Authoritarian Personality (New York: Harper and Row, 1950).

John W. Atkinson, An Introduction to Motivation (New York: Van Nostrand Reinhold Co., 1964).

Reuben M. Baron, "Attitude Change Through Discrepant Action: A Functional Analysis," in A.G. Greenwald, T.C. Brock, and T.M. Ostrom, (Eds.), Psychological Foundations of Attitudes (New York: Academic Press, 1968), 297-326.

Earl R. Carlson, "Attitude Change Through Modification of Cognitive Structure," Journal of Abnormal and Social Psychology, 52 (1956), 256-61.

Arthur R. Cohen, "Experimental Effects of Ego-defense Preference on Interpersonal Relations," Journal of Abnormal and Social Psychology, 52 (1956), 19-27.

Arthur R. Cohen, "Need for Cognition and Order of Communication as Determinants of Opinion Change," in C.I. Hovland (Ed.), The Order of Presentation in Persuasion (New Haven: Yale University Press, 1957), 79-97.

Arthur R. Cohen, Ezra Stotland, and Donald M. Wolfe, "An Experimental Investigation of Need for Cognition," Journal of Abnormal and Social Psychology, 51 (1955), 291-4.

George S. Day, "Theories of Attitude Structure and Change," in S. Ward and T.S. Robertson (Eds.), Consumer Behavior: Theoretical Sources (Englewood Cliffs, N.J.: Prentice-Hall, Inc. 1973), 303-53.

Alice H. Eagly and Samuel Himmelfarb, "Current Trends in Attitude Theory and Research," in S. Himmelfarb and A.H. Eagly (Eds.), Readings in Attitude Change (New York: John Wiley and Sons, Inc. 1974), 594-610.

Samuel Himmelfarb and Alice H. Eagly, "Orientations to the Study of Attitudes and Their Change," in S. Himmelfarb and A.H. Eagly (Eds.), Readings in Attitude Change (New York: John Wiley and Sons, Inc., 1974), 2-49.

Chester A. Insko, Theories of Attitude Change (New York: Appleton-Century-Crofts, 1967).

Daniel Katz, "The Functional Approach to the Study of Attitudes," Public Opinion Quarterly, 24 (1960), 163-204.

Daniel Katz, Charles McClintock and Irving Sarnoff, "The Measurement of Ego Defense as Related to Attitude Change," Journal of Personality, 25 (1957), 465-74.

Daniel Katz, Irving Sarnoff and Charles McClintock, "Ego-Defense and Attitude Change," Human Relations, 9 (1956), 27-45.

Daniel Katz and Ezra Stotland, "A Preliminary Statement of a Theory of Attitude Structure and Change," in S. Koch (Ed.), Psychology: A Study of a Science, Vol. III (New York: McGraw-Hill, 1959), 423-75.

Charles A. Kiesler, Barry E. Collins, and Norman Miller, Attitude Change: A Critical Analysis of Theoretical Approaches (New York: John Wiley and Sons, Inc., 1969).

William B. Locander, and W. Austin Spivey, "A Functional Approach to Attitudes and Consumers' Brand Loyalty," unpublished manuscript, 1975.

Edwin A. Locke, "Personnel Attitudes and Motivation," in M.R. Rosenzweig and L.W. Porter (Eds.), Annual Review of Psychology, Volume 26 (Palo Alto, Calif.: Annual Reviews, Inc., 1975), 457-80.

Richard J. Lutz, "The Functional Approach to Attitudes: A Reconceptualization with Operational Implications," Working Paper No. 53, Center for Marketing Studies, UCLA, July 1977.

Richard J. Lutz, "A Functional Theory Framework for Designing and Pretesting Advertising Themes," in J.C. Maloney and B. Silverman (Eds.), Attitude Research Plays for High Stakes (Chicago: American Marketing Association, 1978), forthcoming.

Charles G. McClintock, "Personality Syndromes and Attitude Change," Journal of Personality, 26 (1958), 479-93.

Henry A. Murray, Explorations in Personality (New York: Oxford Press, 1938).

Helen Peak, "Attitude and Motivation," in M.R. Jones (Ed.), Nebraska Symposium on Motivation (Lincoln: University of Nebraska Press, 1955), 149-89.

- Helen Peak, "The Effect of Aroused Motivation on Attitudes," Journal of Abnormal and Social Psychology, 61 (1960), 463-8.
- Helen Peak, and H. William Morrison, "The Acceptance of Information into Attitude Structure," Journal of Abnormal and Social Psychology, 57 (1958), 127-35.
- Milton Rokeach, Beliefs, Attitudes and Values: A Theory of Organization and Change (San Francisco: Jossey-Bass, Inc., 1968).
- Milton J. Rosenberg, "Cognitive Structure and Attitudinal Affect," Journal of Abnormal and Social Psychology, 53 (1956), 367-72.
- Milton J. Rosenberg, "An Analysis of Affective-Cognitive Consistency," in C.I. Hovland and M.J. Rosenberg (Eds.), Attitude Organization and Change (New Haven: Yale University Press, 1960), 15-64.
- Ivan Ross, "Self Concept and Brand Preference," Journal of Business, 44 (1971), 38-50.
- Julian B. Rotter, Social Learning and Clinical Psychology (New York: Prentice-Hall, Inc. 1954).
- Irving Sarnoff, "Identification with the Aggressor: Some Personality Correlates of Anti-Semitism Among Jews," Journal of Personality, 21 (1951), 199-218.
- Irving, Sarnoff, "Psychoanalytic Theory and Social Attitudes," Public Opinion Quarterly, 24 (1960a), 251-79.
- Irving Sarnoff, "Reaction Formation and Cynicism," Journal of Personality, 28 (1960b), 129-43.
- Irving Sarnoff, "Social Attitudes and the Resolution of Motivational Conflict," in M. Jahoda and N. Warren (Eds.), Attitudes: Selected Readings (Baltimore: Penguin Books, 1966), 279-84.
- Irving Sarnoff, and S.M. Corwin, "Castration Anxiety and the Fear of Death," Journal of Personality, 27 (1959), 374-85.
- Irving Sarnoff and Daniel Katz, "The Motivational Bases of Attitude Change," Journal of Abnormal and Social Psychology, 49 (1954), 115-24.
- M. Brewster Smith, "The Personal Setting of Public Opinions: A Study of Attitudes Toward Russia," Public Opinion Quarterly, 11 (1947), 507-23.
- M. Brewster Smith, "Personal Values as Determinants of a Political Attitude," Journal of Psychology, 28 (1949), 477-86.
- M. Brewster Smith, Jerome S. Bruner, and Robert W. White, Opinions and Personality (New York: John Wiley and Sons, Inc., 1956).
- Ezra Stotland, Daniel Katz, and Martin Patchen, "The Reduction of Prejudice Through the Arousal of Self-Insight," Journal of Personality, 27 (1959), 507-31.
- Morton Wagman, "Attitude Change and Authoritarian Personality," Journal of Psychology, 40 (1955), 3-24.
- Bernard, Weiner, Theories of Motivation: From Mechanism to Cognition (Chicago: Markham, 1972).
- William L. Wilkie, and Edgar A. Pessemier, "Issues in Marketing's Use of Multi-Attribute Models," Journal of Marketing Research, 10 (1973), 428-41.

NEW DIRECTIONS IN ATTITUDE RESEARCH: A CRITICAL EVALUATION

Joel B. Cohen, University of Florida

I have been asked to discuss the four papers presented at the session, "New Directions In Attitude Research." I suppose it would be appropriate to start off with the question, "So what is new?" While I'm not certain any of the authors would claim to have started us down a truly virgin path, consumer researchers can, I think, harken back to four identifiable research orientations which have been given a shot in the arm by these papers. All have been underrepresented in previous consumer research. I have labeled these four major research thrusts as: (1) an information processing approach to attitude formation and change, (2) the importance of time in studying cognitive organization and change, (3) explicit consideration of the motivational bases for attitudes and attitude change, and (4) attitudes as interpersonal communications. I'd like to discuss how each of these has been treated by the authors who have addressed these issues.

Attitudes As Interpersonal Communications

I begin with this topic because it is probably the newest of the attitude research orientations to be represented within the consumer behavior area. The notion of products as identity-relevant symbols, of course, is one of the oldest and best enduring concepts in our field. From the social class and reference group-oriented approaches of the 1950's (cf. Rainwater, et. al., 1959; Levy, 1959) through interpersonal response trait-oriented research (cf. Cohen, 1967; Cohen and Barban, 1970) to self-concept based research (cf. Dolich, 1969; Ross, 1971), consumer researchers have long recognized the importance of identity enhancing social communications to product and brand choice.

Schlenker's "social identity theory" provides a unifying conceptual structure within which to examine self-presentation strategies and behaviors. Attitude expressions and statements are viewed as interpersonal communications--thus, behaviors to be studied in their own right. According to Schlenker, "Even 'private' attitudes, expressed only to oneself, occur in the context of imaginary social settings and serve the same functions in our private worlds." Viewed within the framework of social identity theory, then, a person's attitudes and expressions of attitudes represent a deliberate attempt to create a particular identity, an impression management strategy.

Knowing Barry Schlenker as well as I do, I simply cannot believe that he is expressing favorable attitudes toward his social identity theory merely or primarily to gain respect and admiration or to otherwise stroke his self-concept. No, I am certain he believes very strongly in the worth of this orientation and, in fact, would evidence considerable enthusiasm for it were he isolated on a desert island with no praise in sight. Indeed, were Barry physically restrained from communicating about this topic to any other human being, it would probably play the same guiding and organizing role in his beliefs about human behavior.

Rich Lutz might want us to analyze Barry's attitudes toward social identity theory in terms of a "functional" approach, which I suspect might go heavy on utilitarian and particularly knowledge functions. In fact, what better example of an attitude object based on the knowledge function can be found than a theory in the hands of a dedicated scientist? Such a theory helps the scientist to find meaning out of seemingly disconnected observations, to organize relevant aspects of human behavior and thereby to provide clarity and consistency to his thought processes. We can think of this in terms

of the observations Jerry Olson and Phil Dover have made with respect to the organizing aspects of memory schemata.

The functional approach argues for balance, for a willingness to see attitudes as serving multiple and diverse needs and goals. Social identity theory promises to enrich our understanding of attitudes which, at least in part, serve value expressive needs. Such attitudes, whether expressed in the form of volunteered statements about a particular product, responses to a questionnaire or in overt purchase behavior convey an association between aspects of a desired self-image and certain product characteristics. Under what conditions (e.g., situations, types of products, the nature of the "audience") a consumer's attitudes will be most likely to reflect impression management goals would appear to be a highly relevant, and eminently researchable, topic area.

Motivational Bases For Attitudes

One of the interesting theoretical issues involved in all of the papers is the motivating role of cognitive consistency in human behavior. To what degree is the establishment and maintenance of cognitive consistency a major driving force, goal or function? Rich Lutz's paper addresses this issue by viewing knowledge, cognitive organization and consistency needs as one of the major bases underlying the existence of attitudes in the first place: to provide meaning and structure and as a guide for action. The disruption of cognitive consistency, in this sense, could produce considerable ambiguity, uncertainty and behavioral conflict. Restoration of consistency, such as through exposure to information favorable to one's initial position, should be favorably regarded and conducive to the acceptance of such material. Yet, conflicts frequently exist between one's desire for consistency and one's adjustment needs. Only by being open to new information and being willing to modify existing beliefs as well as behaviors can the individual grow and adapt successfully to a dynamic environment.

Under conditions other than those which threaten to disrupt cognitive systems which the individual is committed to maintain (e.g., central beliefs about one's self, beliefs whose disruption would so disorient the individual as to impact on his ability to act), knowledge and organization-based cognitive consistency processes, while fundamental and pervasive, are probably not dominant.

Sources of Cognitive Inconsistency

Barry's paper makes the point (which other critics have also made) that the concept of cognitive inconsistency as a drive-like state (which then "pushes" dissonance reduction processes) is not well conceptualized or defined by its advocates. Is cognitive inconsistency really all that important? This criticism is, I believe, well founded and stems in large part from the failure of dissonance theorists to come to grips with and spell out both the source and the nature of such drive-like states.

One could, I suppose, attempt to argue that critics have all too often ignored the clearly stated proposition that dissonance theory requires an assessment of the number and importance of dissonant cognitions, and thereby a determination of the magnitude of the dissonance present, before rushing off to make predictions regarding dissonance reduction processes. Clearly,

some cognitive inconsistencies appear to be trivial and others quite important. But, frankly, the theory's advocates have tended to skirt this important issue, and the critics have been left with an inviting (if imprecise and ambiguous) target. Psychological commitment to an action or position, degree of choice, choice incentives and foreseeability, while eminently researchable correlates of the magnitude of dissonance, do not provide a very satisfying conceptualization of the nature and strength of the dissonant state.

One's response to this state of affairs could be to look elsewhere, to turn to other "competing" models in the hope they will be at least as productive and a good deal more precise. Incentive theory (Elms, 1967), self-perception theory (Bem, 1972) and most recently social identity/impression management theory (Tedeschi, et. al., 1971; Schlenker, 1977) represent well-conceived efforts in this direction. While each of these (and particularly social identity theory) can be seen to offer compelling alternative explanations for dissonance-like effects, particularly in the typical counterattitudinal behavior paradigms, I do not find any of these formulations to be as fundamental theoretically as dissonance theory. The potential integrative strength of dissonance theory has not been realized, in part because of a failure to recognize and combine both information processing and self-concept related underpinnings.

What I would like to propose is that there are two major sources of dissonance, in the sense of dissonance being a drive-like state resulting from cognitive inconsistency. The first of these is clearly related to Katz's (1960) "knowledge function" which Lutz has described in his paper. There is a need for meaningful and consistent cognitive organization in order to deal effectively with the external world. Without this, the world would be a mass of disconnected stimuli. In response to such needs, we see the development of cognitive structure, which provides an interpretative and integrative function influencing all information processing activities.

Clearly, there must be an adaptive balance between openness to reality (even if it be counter to our prior beliefs and expectations) and the maintenance of an organized system of cognitive categories (which we take to be our knowledge of what is and what is not, what things go together, etc.). Experiments such as those which confront subjects with a deck of playing cards containing a black queen of hearts or a red king of spades illustrate a principle which has been termed "assimilation to the typical instance" (Jones and Gerard, 1967) and which is of considerable importance for social perception (e.g., stereotyping). Milton Rokeach has emphasized this dimension by defining resistance to change of a belief "solely in terms of connectedness: the more a given belief is functionally connected or in communication with other beliefs, the more implications and consequences it has for other beliefs and, therefore, the more central the belief" (1968, p. 5).

Very few dissonance studies have specifically examined the drive-like effects of a disruption in one of a person's central beliefs. It is not an easy matter to design a study which would seriously challenge an important personal belief. There can be no doubt, however, that the dissonance caused by strongly believing A implies B and then learning that A implies not B would be drive-like: some resolution of this cognitive dilemma must be found to restore order and meaningfulness. (See Abelson, 1959 for a discussion of denial, bolstering, differentiation and transcendence in resolving belief dilemmas.) Carried to an extreme, "It may be supposed that any inexplicable disruption of these taken-for-granted constancies, physical or social or

self, would lead one to question the validity of one's own senses, one's competency as a person who can cope with reality, or even one's sanity" (Rokeach, 1968, page 7). Field studies which have the potential to create powerful disconfirmation of beliefs (supported by behavioral commitment) as when one who is committed to a course of action is suddenly and dramatically confronted by evidence that his beliefs and course of action are wrong (see for example Festinger, et. al., 1956) may provide a more appropriate setting in which to examine the power of cognitive inconsistencies than the usual types of scenarios presented to subjects in laboratory studies.

The major source of dissonance is probably inconsistencies relating to the self-concept. Here the issue is not coping with reality and cognitive organization, per se, but rather what Katz (1960) has termed "value expressive" needs. It is quite true, of course, that self-concept related beliefs may be important anchoring points for a large number of related beliefs, and hence play an important organizing function. Such beliefs should be quite resistant to change, and behavior seen as inconsistent with them should provoke considerable dissonance. It is with self-concept related beliefs that dissonance theory and social identity theory may be inextricably linked. Whether or not social identity theory is a competitor or serves to extend and clarify may be more a matter of judgment than an empirical question.

Another reaction may be that what is proposed here is no longer quite dissonance theory or at least is a substantial modification of it (perhaps extending Aronson's, 1968, line of thinking). The basic point, though, is that the theory itself is unusually silent on such a key issue as the nature and identification of the "importance" of a cognitive element. This was acknowledged as early as 1962 (Deutsch, et. al.) when the following proposition was advanced: "Thus, we offer the hypothesis that a chooser will experience post-decisional dissonance only when he perceives his choice in a given situation to be inconsistent with the conception of some aspect of himself which he tries to maintain (for himself or for others) in that situation."

Dissonance Reduction Versus Self Enhancement

What then is the motivating force, inconsistency among cognitive elements one or both of which serve important self-defining or organizational functions, or (per Schlenker) a desire to act so as to increase personal association with desirable attributes and outcomes and to decrease such association with undesirable attributes and outcomes? Social identity theory sees the degree of fit between the identity the individual is trying to project and some action, attitude, or choice object as the key element. Dissonance theory would conceptualize this fit or association in terms of cognitions of varying importance and consistency.

Paired in this way, the commonality among these formulations may be more important than the differences. Social identity theory greatly extends, clarifies and refines the nature of attitude-behavior relationships which involve social identities and related behavior.

Schlenker and Schlenker (1975, p. 963) state, "The greater the degree of personal responsibility an individual perceives for an action which generates negative consequences, the greater his degree of personal dissatisfaction, anxiety, and attendant social punishment." The dissonance theory position proposed earlier leads to a very similar hypothesis, since negative consequences of one's behavior should be strongly inconsistent with a very important set of cognitions: those which define and proclaim one's self-concept. The Schlenkers, in an

insightful overview of the counterattitudinal behavior paradigm so often used in dissonance and self-perception theory research, point out that these "have placed individuals in positions where their self-image and public image is jeopardized..." (p. 964). While they maintain that "attitudinal discrepancy is not a necessary antecedent to attitude change," I would argue that the discrepancy between self-concept attitudes and inconsistent behavior is precisely the nature of the individual's predicament. Social identity theory, to its credit, goes on to identify tactics to be used by the individual to maintain and enhance his self-image and presentation of self. These same tactics, of course, serve to provide the individual with cognitions regarding his behavior which are more consistent with his self-concept.

The conceptualization of self-image maintenance and enhancement tactics appears to be quite similar to that adopted by Lutz in his treatment and proposed operationalization of the functional approach. By positioning the functional approach within an expectancy-value framework, Lutz develops the similar (to Schlenker) notion that attitude change attempts (whether externally or internally generated) must effect either the perceived association (i.e., perceived instrumentality) between the individual and the action (e.g., taking personal responsibility in Schlenker's system) or the value of the consequences (e.g., minimizing the amount or existence of harm in Schlenker's system). Perception of choice manipulations are a favorite device used to impact on the amount of responsibility (perceived instrumentality) the experimental subject perceives he had for whatever consequences developed. When personal responsibility is heightened, and the individual is therefore unable to excuse or justify his actions, he will frequently attempt to alter his perception of the resulting consequences (e.g., "No 'real' harm occurred; At the time I said, wrote or did X I really believed it was the good, true, or the right thing to do").

The same type of expectancy-value analysis has been applied by Schlenker and Schlenker (1975) to the situation in which positive rather than negative consequences result. Specifically, an average to dull person was told that she was quite attractive under either high or low choice (responsibility) conditions. S's subsequent scores were more positive than those of a control group only when there was low choice and S's expected future interaction during the study with the other person. A social identity theory expectancy-value analysis explains this result quite nicely in terms of S's exhibiting favorable attitudes toward the other person in order to associate themselves with and claim responsibility and credit for the positive consequences stemming from their (experimenter imposed) communications to the other person.

Schlenker and Schlenker feel that attitude change, which has the effect of establishing an internal locus of responsibility, is counter to dissonance theory predictions (i.e., S's having no choice and producing positive consequences should, they claim, experience no dissonance).

The problem, however, is that dissonance theory is ambiguous in terms of the source of the cognitive inconsistency. If, as the Schlenkers acknowledge, the dissonant cognition may relate to the self-concept (e.g., "I'm a friendly and helpful person. Yet I'm coming across as if I might not be that way at all because the experimenter wouldn't let me exhibit these qualities-- I was given no choice."), then dissonance reduction is a viable alternative explanation. The Schlenkers' unwillingness to accept this explanation seems to be more a matter of overall discontent with dissonance

theory's lack of precision (or an inability to pin it down, if you like) than the responsiveness of this particular explanation to their data.

Actually, there may be an even simpler dissonance theory interpretation: S's were merely bringing their questionnaire responses into line with their anticipated behavior. The no-choice, future interaction S's probably anticipated behaving in a positive and favorable manner toward the other experimental subject when the experimenter brings them together. Such S's will also know that the experimenter has told the other person that the favorable communication was assigned rather than chosen. These S's should have a greater reason to show both the experimenter and the other person that their attitudes and subsequent behavior are, in fact, consistent than S's whose free choice has already conveyed this.

Now, whether this preference for consistency stems primarily from a desire for cognitive order (i.e., to eliminate confusions that might interfere with one's ability to behave clearly and convincingly-- see Jones and Garard, 1967, for a discussion of the value of an "unequivocal behavioral orientation."), or a desire to appear consistent and genuine (to be otherwise would be inconsistent with important self-concept attitudes) can't be definitively resolved in this study. Even worse, it looks like it is going to be very difficult indeed to design a study in which dissonance reduction and the maintenance and enhancement of social identity (as alternative explanations) are not somehow confounded.

Concerning the need for jousting with dissonance theory, my recommendation to Barry would be to let the dragon slumber. It's been pretty quiet of late and doesn't seem to be up and about too often. There isn't really any need to take it on. The issues, settings, and most importantly the interpersonal goals which are directly addressed by social identity theory are not directly captured or represented by the more cognitively oriented dissonance theory.

Thus, even though dissonance theory may provide a somewhat more elegant and over-arching theoretical position which can incorporate virtually any set of data or conceptualization which possesses consistency-striving characteristics (whether to an actual or ideal self-image or presentation), it offers comparatively few insights into the self presentation and impression management areas when compared to social identity theory. While theoretical validity, comprehensiveness and parsimony may be ultimately satisfying to the soul of any behavioral scientist, the usefulness, adequacy and operational vitality of a conceptual framework may be a good substitute for the working social psychologist. On this score, and given the area on which Barry has chosen to focus, social identity theory appears to be off to a flying start. That one person or another may explain the same percentage of variance in one's data by some simple or "oblique" rotation of conceptual axes in and of itself does not establish the "truth" of that conceptual system.

Social identity theory is at one and the same time both less and a good deal more ambitious than any generalized consistency theory. It is less ambitious because it presents no claims on areas of inquiry which do not reside within its narrower domain (e.g., Olson and Dover's concern with information processing aspects of cognitive consistency). It is more ambitious because it does seek to provide a much more adequate and complete description and explanation of the many facets of human behavior concerned with social identity and self-presentation. This is a difficult and challenging area and one of great relevance to social psychologists and consumer behaviorists alike. There is no doubt that

we shall be hearing a great deal more about social identity theory in the near future.

Status of the Functional Approach

I'd like to turn now to a consideration of Rich Lutz's excellent review and analysis of the so-called "functional approach to the study of attitudes." There are a number of reasons for advocating a categorization scheme similar to the one Rich has been working on. First, of course, such a scheme has the potential for shaking us out of any intellectual lethargy we may have fallen into in thinking too narrowly about such things as the antecedents, support for and conditions for changing attitudes and behaviors. There is also a specificity problem inherent in theories conceptualized at a highly abstract level (e.g., learning theory, dissonance theory), especially when considering applications to an area like consumer behavior, although such theories are generally considered to be more fundamental as well as parsimonious.

The general notion of differing functional bases seems sound enough. Attitudes and behavior expressive of prejudice toward a minority group, or for that matter, of liking for a dashing new sports car, may have very different functional bases and goals. Lutz's extension and modification of Katz's (1960) categorization scheme for the consumer behavior area may well help us to keep this fact firmly in our minds.

Another reason for advocating such a categorization scheme is as a kind of motivational check list (e.g., before leaving on a long trip, stop the newspapers, call the post office, lock the doors and windows, etc.). It is here that Rich may run into some trouble. At what level of specificity should functions be stated so as to be useful, say, to those interested in changing existing attitudes? My check list of things to do before leaving on a long trip is quite specific and at an actionable level. In addition, I have some confidence that these are, in fact, the things to do and that there aren't any major omissions. Can the same be said of Lutz's listings? Kiesler, Collins and Miller (1969) had earlier leveled this type of criticism at the functional approach.

If what we are after is a check list which is both more complete and more precise, we might begin with a review of the literature pertaining to personality needs and goals, from the pioneering efforts of McDougall (1908) through Murray (1938) to the recent day factor theories (see, for example, Cattell and Scheier, 1958). Clearly, the number and diversity of human needs, motives and goals can approach staggering proportions if our goal is to be exhaustive or purely descriptive (as in psychographic research).

It seems clear that Rich must be striving for something different than an exhaustive check list, some midpoint at which the structure provided attains that happy balance between pointing and naming, on the one hand, and abstract conceptualization on the other. While researchers have never been able to reach a consensus on how many functions, or needs or goals is the "right" number, not to mention what these are, that probably should not foreclose efforts at developing the sort of categorization which Rich believes would be useful. It seems clear, however, that "merely" developing satisfactory operational measures for the functions he advocates does not address the more basic question regarding the adequacy of the conceptualization of needs and goals represented by the classification scheme.

In summary, it appears as though the jury is still out, and a verdict cannot yet be given as to the usefulness of the functional approach in consumer behavior. If it

does nothing more than alert us to the need to look beyond utilitarian product attitude in assessing consequences of using (or not using) a particular product, it will have contributed to our understanding of consumer attitudes and attitude change. We will look forward to greater operational specificity and empirical research on the part of those who advocate this approach.

Information Processing Aspects

The Olson and Dover paper is important for consumer researchers particularly because it raises a set of issues which cannot be ignored by those who would focus on the outcome of information processing and decision making activities extending over several days, weeks or even months.

Changes in Cognitive Structure Over Time

Time (that bane of a researcher's existence, because it forces us out of one-shot measurement paradigms) is the catalyst. It is inconceivable that we will be able or willing to generalize our findings regarding information processing and decision-making very far without adequately investigating changes in cognitions and cognitive organization that occur over time.

Olson and Dover go on to analyze the impact of time on cognitive structure by looking at data from a longitudinal study designed to investigate disconfirmation of expectations. Obviously, some compromise must be made whenever the data one uses are not exactly what one would have gathered for the present purpose, and the authors acknowledge this limitation.

A major result, which Olson and Dover seek to understand and explain, is the increasing consistency among elements of cognitive structure over time. One suggestion that could be made with respect to research of this kind is that cognitive structure is a property of an individual's cognitive system, and hence the degree of consistency should be studied on an individual rather than group level. I'm more than a little uncomfortable with group means and correlations as indicants of changes in an individual's cognitive structure over time.

Evaluative and Probabilistic Consistency

An important distinction with respect to inference processes has been highlighted by the authors who call attention to the parallel concepts of evaluative and probabilistic consistency. The latter refers to the degree to which two concepts seem to be logically related to one another based on a combination of learning and stimulus cues so that inferences regarding an object's possession of other attributes follow along the lines of a conditional probability statement. Inferences based on evaluative consistency, on the other hand, seem not to be as concept-driven as they are affect-driven. That is, I may attribute positive qualities such as intelligence and warmth to a person because he is portrayed positively, not because I believe there is an associative link between the two attributes. Aspects of this process have been referred to as a "halo effect." These, then, are alternative mechanisms for filling in the gaps in one's belief systems, and since all information acquisition is necessarily incomplete and leaves gaps, these are very important processes.

Olson and Dover seek to disentangle the results of these two processes by looking at the increasing consistency they observe among $b_i e_i$ elements. I would recommend looking at the b_i elements by themselves since the concept of evaluative consistency pertains to the basis for the inference process (i.e., a positively regarded object is thought to possess positive attributes), the result of the inference being the

association of the object with some attribute. Once again, though, it is hard to see how this can be done on anything other than an individual basis. In addition, disentangling these processes will require a study specifically designed for this purpose. While the authors suggest that evaluative consistency is operating in their study, this appears to be highly speculative in the absence of data specifically establishing the probabilistic linkage among the belief elements (i.e., the probability of A given B) as well as an independent evaluation of attributes.

Their group data (Figure 1) reveals wide fluctuation in attribute evaluations until people first tasted the bitter coffee, at which point four of the five attribute evaluations moved in the same direction (i.e., became more negative). Then, after an extended trial of the unadulterated coffee, four of the five attribute evaluations became more positive. The authors, looking at intercorrelations among attribute evaluations at the end of the study, concluded that the three attributes most closely related to overall product attitudes were also closely related logically (i.e., probabilistically). While these findings are quite interesting, the data used is simply not adequate to the delineation of inference mechanisms. It is simply not sufficient to estimate the degree of internal consistency among attribute evaluation scores. This sheds no light on the process that led to such a level of consistency.

Separating Out Sources of Impact

One potentially troublesome research issue suggested by the use of multiple measurements in a within-subjects longitudinal design is that our attribute-specific measures may be consistency-inducing. Wyer (1976) has referred to this not uncommon phenomenon as the "Socratic effect." That is, we may be making a person more aware of the probable existence of an attribute-based belief structure and its presumed "logical" relationship to one's overall attitude. In so doing, we may encourage him to think through any inconsistencies in his own belief structure and, particularly for insightful experimental S's, to resolve these in the direction that would allow us to perceive them as more logical.

The authors make the important observation that beliefs may not be well-formed or stable after an initial exposure to information about an object or even the object itself: it may take several such experiences with the object. They argue for the direct measurement of S's confidence in their belief ratings. Here again, a within-subjects design may cause some problems. It may be that S's will believe that they should become more confident as a result of increased information and/or experience.

Olson and Dover report that the evaluation associated with beliefs regarding bitterness was, by the end of the study, the most strongly related attribute to overall attitude. This may, of course, simply be the result of demand effects since there were three separate communications designed to convince people that the coffee was not bitter followed by a "bitter" manipulation, and S's may have gotten the idea that bitterness was somehow important, at least to the researchers.

It is interesting, however, that the correlation between the evaluation of this attribute and overall evaluation was not particularly strong nor showing any particular trend until people actually experienced the bitter coffee. Perhaps this is still another illustration of the difference in impact between communications as information and experience as information. It may be something to think about in designing experimental

manipulations, not to mention the various reasons to be interested in attitude and behavior change strategies! Perhaps one explanation for the greater impact, following the thrust of Olson and Dover, may be the greater "depth of processing" that is probably associated with the increased involvement overt behavior typically generates (See Nisbett, et. al., 1976, for a related viewpoint).

This particular piece of data also reminded me of the unique problems (e.g., long hours spent trying to communicate to respondents followed by blank stares) involved in the association between the absence of a negative state (i.e., this coffee is not bitter) and one's resulting positive evaluation. This double negative leading to a positive may not be as easily learned as a positive state leading to a positive evaluation.

There are, after all, a very large number of attributes an object doesn't possess, most of which we are not aware of to begin with or ignore. Saying that an object doesn't have something may actually communicate less information than saying it does have something, when the attribute in question is not highly salient to begin with. Thus, communications which instruct me that a particular brand of coffee does not possess an attribute I hadn't given much thought to will not produce very much cognitive work: there is nothing more I must deal with. By comparison, instructing me that the coffee does possess an attribute I hadn't given much thought to provides me with information which I need to integrate with existing information. The informational act of tasting the bitter coffee and experiencing its consequences should have given people a much better reason to link the absence of that attribute with their overall evaluation of the coffee.

Time and the Decay of Advertising Effects

Olson and Dover have enlightened us as to the substantial changes which may take place in cognitive structure over time, and Sawyer and Semenik, likewise, stress the need for researchers to examine delayed as well as immediate effects of advertising. The major thrust of the paper has to do with the decay of corrective advertising effects. Sawyer and Semenik invoke Jost's Second Law of Learning to explain why a corrective ad may appear to be more effective in reversing the effects of a deceptive campaign than it really is. According to Sawyer and Semenik, this translates to mean that "the effect of the older of the two (stimuli) will decay less rapidly - reaching some asymptotic level higher than that of the newer campaign."

The rate of decay (i.e., memory based) argument is quite similar to the explanation frequently given for observed recency effects when dependent measures are administered shortly after the second communication and some interval of time has elapsed between the first and second communications.

I have some difficulty with the higher asymptotic level of the older stimulus without knowing a great deal more about the reward and practice schedules which might so forestall further decay as to give it absolute superiority over a newer stimulus. Phrased more colloquially, Sawyer and Semenik's Figure 2A suggests to me that the older stimulus is comparable to a George Allen established performer and newer stimulus to an unproven rookie. A specific a priori determination of relative belief strength toward to the two stimuli, in my opinion, will not be derivable from any law of learning which fails to specify the exact relationships between each stimulus and other pertinent

beliefs together with the value of the respective items of information to the individual. This is a tall order!

Another compelling interpretation may be that processing of the information in the second communication may be incomplete if measures are administered immediately thereafter, with the result that a "play back" of the position advocated in this communication (assuming it has been presented in a convincing manner) is a "sensible" response to the questions we are asking. With continued processing of the message, the individual should have a much greater opportunity to compare both its conclusions and premises with related stored information. If aspects of the new information are not as consistent with previously well-supported beliefs as it appeared on first contact, it may then be treated with less than perfect certainty or otherwise qualified. In any event, both memory-based and information processing-based arguments support the taking of delayed measures.

Time and Cue-Dissociation

I have related problems with the "discounting cue hypothesis" which assumes that a dissociation of an inhibiting cue (the authors look upon a corrective ad as an inhibiting cue) over time will reinstate the effects of the original stimulus. First, I have difficulty with translating a sometimes empirical observation into an explanation. Some information had less effect over time. Why? Because people discounted it. Very enlightening!

Perhaps a closer look should be taken at the authors' descriptions of the characteristics of information that was apparently "discounted" in other research. This might provide some idea as to why people might behave in this fashion. Let me throw out just one possibility. Whereas the non-discounted information seems to combine both conclusions and supporting arguments, the discounted information appears only to have offered a summary blanket refutation. Unless this refutation succeeded in removing the previous information from memory storage, further reality and consistency testing of that information was possible, indeed likely because of the conflicting conclusions which had been transmitted. No such opportunity was apparently present for arguments underlying the refutational conclusion, because none were presented. So a much greater opportunity seems clearly present, then, for people to make a subsequent assessment as to whether the "positive" conclusion "rings true."

Additionally, I'm not sure that Sawyer and Semenik have made a very good case for applying such concepts directly to the corrective advertising area. As I tried to indicate above, we need to take a long and hard look at both the concepts advocated and how they are operationalized. A concept invariably contains surplus meaning beyond a given operationalization. Perhaps we ought to be a little more conservative and claim no more territory than we have empirically staked out. In a sense, this is another argument for the importance of construct validity in behavioral research. To the extent that the corrective advertisement operationalizations are quite different from the research paradigms on which the concepts are based, a healthy skepticism seems in order. Identification of the major differences between the two contexts and settings may be a necessary and important step.

Let's just take one more brief example which spotlights the nature of the problem. Sawyer and Semenik present some findings indicating that qualified messages (stated more equivocally and with more reservations) persist better than unqualified messages. One set of

now familiar explanations is that the qualifications and reservations are forgotten, discounted or dissociated from the main conclusions over time.

But what if this effect is due instead to the greater perceived credibility, expertise and trustfulness of a communicator who appears to be trying to be fair to both sides as well as more objective? It certainly looks like that person has been very thorough and is clearly well informed (more so that I could be on the topic). I might tend to believe his conclusions to a far greater extent than those of the communicator who has presented an unqualified message. Sawyer and Semenik seek to equate qualified messages with the format of corrective advertisements. But, if the communicator credibility notion is correct, and we assume that the credibility of an FTC-initiated conclusion is quite high, making the conclusion more equivocal and tentative may not only be unnecessary but a serious mistake.

Unfortunately, the ambitious study designed by Sawyer and Semenik appears to have misfired with respect to the stimulus materials themselves (i.e., there were no significant differences on the target belief between a control group of people who were not exposed to any message and S's exposed to either of the experimental messages). It would not appear to be particularly useful, therefore, to attempt to interpret some of the other results (a number of which are non-obvious) in the light of the difficulties and suggestions for improvement the authors have themselves raised in the paper.

References

- E. Aronson, "Dissonance Theory: Progress and Problems," in R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg and P. H. Tannenbaum (Eds.), Theories of Cognitive Consistency: A Sourcebook, (Chicago: Rand McNally), 1968.
- D. J. Bem, "Self-Perception Theory," in L. Berkowitz (Ed.), Advances in Experimental Social Psychology (Vol. 6), (New York: Academic Press), 1972.
- R. B. Cattell and I. H. Scheier, "The Nature of Anxiety: A Review of Thirteen Multivariate Analyses Comprising 816 Variables," Psychological Reports, 4(1958), p. 351-388.
- J. B. Cohen, "An Interpersonal Orientation to the Study of Consumer Behavior," Journal of Marketing Research, 4(Aug., 1967) p. 270-278.
- J. B. Cohen and A. Barban, "An Interactive Consumer-Product Typological System: A Progress Report and Partial Evaluation," Working Series in Marketing Research, Number 12 (College of Business Administration, The Pennsylvania State University), 1970.
- M. Deutsch, R. M. Krauss, and N. Rosenau, "Dissonance or Defensiveness," Journal of Personality, 30(March, 1962), p. 16-28.
- I. J. Dolich, "Congruence Relationships Between Self-Images and Product Brands," Journal of Marketing Research, 6(Feb. 1969), p. 80-84.
- A. C. Elms, "Role-playing, Incentive, and Dissonance," Psychological Bulletin, 68(1967), p. 132-148.
- L. Festinger, H. Riecken, and S. Schachter, When Prophecy Fails. (Minneapolis: University of Minnesota Press), 1956.

- E. E. Jones and H. B. Gerard, Foundations of Social Psychology, (New York: Wiley), 1967.
- D. Katz, "The Functional Approach to the Study of Attitudes." Public Opinion Quarterly, 24(1960), p. 163-204.
- C. A. Kiesler, B. E. Collins and N. Miller, Attitude Change, (New York: Wiley), 1969.
- S. J. Levy, "Symbols for Sale," Harvard Business Review, 37(July-August, 1959), p. 117-124.
- W. McDougall, An Introduction to Social Psychology, (London: Methuen and Company), 1908.
- H. A. Murray, Explorations in Personality, (New York: Oxford), 1938.
- R. E. Nisbett, E. Borgida, R. Crandall and H. Reed, "Popular Induction: Information is not Necessarily Informative," in J. S. Carroll and J. W. Payne (Eds.), Cognition and Social Behavior, (Hillsdale, N. J.: Lawrence Erlbaum Assoc.), 1976.
- L. Rainwater, R. P. Coleman and G. Handel, Workingman's Wife, (New York: Oceana Press), 1959.
- M. Rokeach, Beliefs, Attitudes and Values, (San Francisco: Jossey-Bass), 1968.
- I. Ross, "Self Concept and Brand Preference," Journal of Business, 44(Jan. 1971), p. 38-50.
- B. R. Schlenker and P. A. Schlenker, "Reactions Following Counterattitudinal Behavior Which Produces Positive Consequences," Journal of Personality and Social Psychology, 31(1975), 962-971.
- B. R. Schlenker, "Translating Actions to Attitudes," unpublished research proposal (University of Florida, Department of Psychology), 1977.
- J. T. Tedeschi, B. R. Schlenker and T. V. Bonoma, "Cognitive Dissonance: Private Ratiocination or Public Spectacle?" American Psychologist, 26(1971), p. 685-695.
- R. S. Wyer, Jr., Cognitive Organization and Change: An Information Processing Approach, (Potomac, Md.: Lawrence Erlbaum Assoc.), 1974.

SHORT AND LONG TERM COST CONSEQUENCES OF GOVERNMENT REGULATIONS TO CONSUMERS

Rachel Dardis, University of Maryland

Abstract

Government regulations in the area of textiles are categorized according to whether they are designed to protect the worker, the environment or the consumer. In most instances regulations will entail an increase in manufacturing costs which will be passed on to consumers in the long run. However, the existence of imports may serve to depress price increases in the case of regulations affecting the worker or the environment.

Introduction

It is not possible in this paper to cover all government regulations affecting the textile industry. However, most government regulations may be classified under three major headings: (a) regulations to protect the textile worker, (b) regulations to protect the environment, and (c) regulations to protect the consumer. Major regulations in each of these areas will first be discussed while the impact of such regulations on the industry and the consumer will be discussed in the second section of the paper.

Regulations to Protect the Textile Worker

Two major regulations address the exposure of the worker to cotton dust and noise. Cotton dust is believed to cause byssinosis though the specific agent in cotton dust has not yet been identified. The current OSHA regulations for cotton dust exposure are set at 1 mg/M3 of total dust (ATMI, 1975). The standard also requires that administrative or engineering controls to reduce cotton dust be implemented where feasible. If such controls are not feasible then respirators or other protection equipment must be worn by the worker. OSHA has proposed a new standard for cotton dust of 0.2 mg/M3 of total dust (Textile World, 1977). Estimated costs of meeting the standards, assuming that equipment could be developed, include initial capital expenditures of at least \$860 million, and annual energy and maintenance costs of \$35 million and \$4 million respectively (ATMI, 1975; Textile World, 1977). The feasibility of the proposed standard was also challenged by a cotton-dust expert at a recent hearing. Hocutt claimed that "anything under 0.5 would shut down spinning; anything under 1.0, weaving" (Textile World, 1977, p. 23).

Noise regulations affect most industries including the textile industry. However, they have a significant impact on the textile industry in view of the fact that textile processing is characterized by relatively high levels of noise. The present OSHA regulations specify that employees shall not be exposed to noise in excess of 90 dBA for an eight hour time period (ATMI, 1975). If noise levels are exceeded then engineering controls must be implemented or protective devices such as ear plugs and muffs must be worn by the worker. However, protective equipment is viewed as a temporary device by OSHA since there is no guarantee that it will be used by the worker. Recently the EPA has proposed an 85 dBA standard which would entail even greater costs for the textile industry. A study by OSHA estimated that it would cost \$1.1 billion to meet the present 90 dBA level and \$2.7 billion to meet the 85 dBA level (Storey, 1976).

Regulations to Protect the Environment

A major regulation in this area pertains to water pollution. Regulations for controlling the discharges of pollutants by the textile industry were promulgated in Feb-

ruary 1974 (ATMI 1975). Two levels of abatement were proposed with the lower level recommended for 1977 and the higher level for 1983. According to industry spokesmen there are few problems in meeting the 1977 level (ATMI, 1975). However, the feasibility of meeting the 1983 level has been debated. Estimated costs to the textile industry in achieving the 1977 and 1983 levels were estimated by the National Commission on Water Quality as follows:

(\$ million)	1977	1983	Total
Capital required	384-780	526-785	910-1,565
Annual Operation & Maintenance	49-68	50-81	

Regulations to Protect the Consumer

The Consumer Product Safety Commission began operation in 1973 after passage of the Consumer Product Safety Act. Four major areas of enforcement are (a) clothing textiles, (b) carpets and rugs, (c) children's sleepwear, and (d) mattresses (Lyons, 1976). Standards have also been proposed for upholstered fabrics and furniture and for wearing apparel.

Studies of the existing children's sleepwear standards concluded that they would increase consumer costs by approximately 30% (Dardis, 1977). However, the standards were also cost effective in view of the number of flammable fabric incidents involving children's sleepwear. Extension of existing sleepwear standards to children's clothing resulted in unfavorable cost-benefit ratios even under the most favorable assumption that the standards would entail no reduction in consumer choice. Cost-benefit analysis has not been applied to either of the two proposed standards. The proposed upholstery standard is estimated to result in a 10% to 30% price increase at retail as well as the elimination of certain fabrics (Storey, 1976). The proposed general wearing apparel standard combines flammability characteristics of textiles with garment style and is designed to reduce flammability treatment costs since tight fitting garments have less severe fabric flammability requirements. However, this standard may also entail a certain reduction in consumer choice since certain fabric/garment combinations will be eliminated. Testing and record keeping requirements for manufacturers will also entail costs.

Impact on Consumers

Impact on U. S. Textile Industry

In assessing the impact of government regulations on the consumer it is first necessary to take into consideration the impact of such regulations on the textile industry. The textile industry is characterized by perfect competition or monopolistic competition, both of which are competitive market structures. In addition the industry faces competition from imports from low wage countries such as Korea, Taiwan and more recently Red China.

Recognition of the textile industry's vulnerability to imports and the need for maintaining a viable domestic industry has led to trade regulations to control the flow of imports into the U. S. The most recent agreement was the MultiFiber Agreement (MFA) which is due to expire at the end of 1977. In spite of this agreement imports continue to present a problem and there has been increased

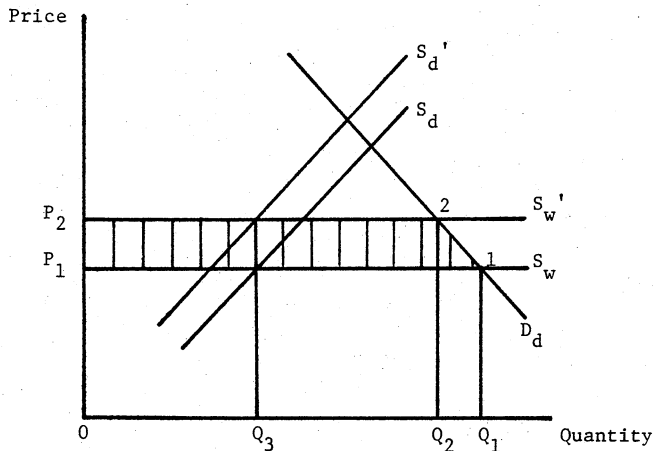
awareness of the necessity for modernization in order to meet competition from other countries. However, the textile industry's ability to modernize may be constrained by the diversion of funds from productive investment to investment to meet government regulations in the area of worker or environmental protection. According to Storey "for the abatement of cotton dust, noise and water pollution, there is a capital requirement equal to five years of the industry's earnings and equal to five years of its average annual investment in new plant and equipment" (Storey, 1976, p. 20). In addition operating and maintenance costs of meeting government regulations in these areas will reduce after tax profits of the industry and hence its earnings for capital expenditures. The reduction in productive investment by the textile industry will mean higher prices for the consumer in the long run.

Short-Term Cost Consequences to Consumers

The impact of government regulations on consumers is illustrated in Figure 1.

FIGURE 1

Cost of Government Regulation



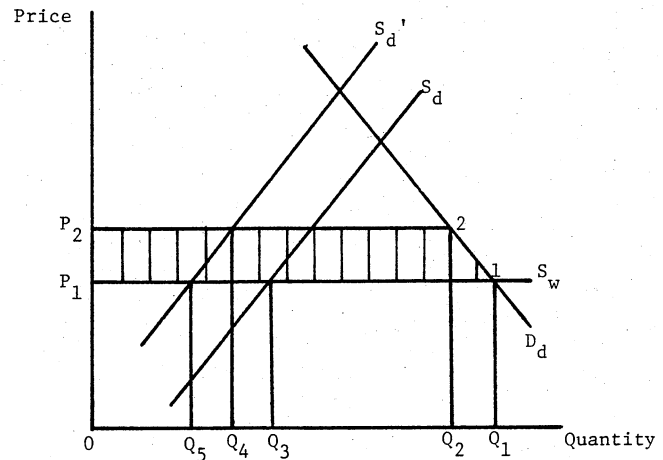
In the initial pre-regulation situation the domestic demand and supply curves are given by S_d and D_d respectively while the world supply curve is given by S_w . A market price and quantity of P_1, Q_1 is obtained with domestic production accounting for Q_3Q_1 units. Imposition of a governmental regulation will increase manufacturing costs by P_2P_1 , corresponding to a shift in the supply curve from S_d to S_d' . The world supply curve S_w may shift or remain constant depending on the nature of the regulations and the policies of exporting countries.

In the case of flammability standards, such standards must also be met by foreign manufacturers. Assuming foreign manufacturers face similar cost increases to those met by domestic manufacturers then S_w will shift to S_w' . The new equilibrium price and quantity is P_2, Q_2 . The cost to the consumer is given by the shaded area P_2P_112 which represents the change in consumer surplus due to the price increase (Dardis, 1977).

In the case of worker or environmental regulations there is no change in the world supply curve unless exporting countries adopt similar regulations. This is unlikely since most textile exporting countries are underdeveloped countries with a need for trade and foreign capital which can only be obtained if they maintain their competitive position in the world market. The impact of government regulation (other than flammability standards) is shown in Figure 2.

FIGURE 2

Cost of Government Regulation



The initial equilibrium situation of P_1, Q_1 is identical to that shown in Figure 1. Imposition of a regulation results in a shift of the supply curve from S_d to S_d' while S_w remains unchanged. Two situations are now possible. S_w In the first instance imports expand to meet consumer demand increasing from Q_3Q_1 to Q_5Q_1 while domestic production declines from Q_3Q_1 to Q_5Q_1 . This situation is unlikely, however, since the flow of imports is controlled by trade regulations. In the event that trade regulations are continued, imports are maintained at their former level (Q_3Q_1) while domestic production declines from Q_3Q_1 to Q_5Q_1 . The final equilibrium price and quantity is given by P_2, Q_2 and the cost to the consumer from the price increase is given by the shaded area, P_2P_112 .

A comparison of Figures 1 and 2 indicates some interesting differences in the impact of government regulations on the consumer. While most regulations will entail a cost increase for manufacturers the degree to which the cost increase is passed on to the consumer depends on the particular regulation. In the case of flammability standards both foreign and domestic manufacturers must meet the standard so that the consumer will pay the full cost increase. In some instances the foreign manufacturer may refrain from exporting to the U. S. in view of the availability of other markets and the cost of operating two production lines. In this case the equilibrium price and quantity after the passage of a flammability standard is given by the intersection of the domestic demand and supply curves (S_d' and D_d). Consumer protection, which discourages imports, also ensures producer protection. In contrast to this result, other regulations place the domestic manufacturer at a competitive disadvantage since they cannot be imposed on foreign manufacturers. Unless trade regulations are in effect the domestic industry will suffer disruption. Even with trade regulations the availability of imports means that the full cost increase is unlikely to be passed on to the consumer. Thus the price increase in Figure 2 is less than the cost increase.

Long-Term Cost Consequences to Consumers

Such costs are more difficult to project. Possible long-term cost consequences include the following:

- (a) exit of firms in the domestic textile industry with a consequent reduction in competition,
- (b) change from cotton textile production to man-made textile production due to uncertainty about cot-

- ton dust standards and flammability standards, which would mean a significant reduction in consumer choice, and
- (c) increased energy requirements by the industry to meet government regulations and a reduction in the supply of energy available to consumers.

Conclusion

As mentioned earlier only a selected number of government regulations could be reviewed in this paper. In particular the impact of the Toxic Substances Control Act was not discussed. It is obvious, however, that the textile industry has faced and will continue to face numerous regulations which are promulgated and enforced by many different government agencies. Several conclusions may be drawn from a review of existing regulations:

- (a) Regulations are needed to protect the worker, the environment, and the consumer. Thus in the absence of regulations (or other measures such as taxes) there is no incentive for the individual firm to control dust or noise levels or water pollution since such controls would increase manufacturing costs and place the firm at a competitive disadvantage.
- (b) Prior to the imposition of the regulations the textile products consumer was not paying the true cost of production. The firm's costs reflected the private costs of production but did not include the social costs of worker ill-health or environmental pollution.
- (c) The implementation of regulations means in most instances an increase in manufacturing costs and such costs will eventually be born by the consumer.
- (d) Flammability standards differ from other regulations since the potential beneficiary also incurs the costs. The more specific the regulation, i.e. addressed to certain high risk groups, the greater the correspondence between costs and benefits.
- (e) Much of the current debate concerns the degree of protection which should be required and the necessity for equating costs and benefits from regulations. Thus in the case of water pollution the 1985 goal of zero discharge is increasingly being questioned on the basis of incremental costs and benefits.
- (f) The increase in the number of government regulations affecting the textile industry requires increased coordination on the part of various government agencies if such regulations are to be implemented in the most cost effective manner.

References

American Textile Manufacturers Institute, "Background on Selected Areas of Regulatory Activity Affecting the U. S. Textile Industry", December 1975.

"Can Textiles Meet Proposed Dust Standards?", Textile World, June 1977, p. 23.

R. Dardis, "Cost-Benefit Analysis of Consumer Product Safety Programs", Final Report, NSF-RANN, 1977.

D. W. Lyons and H. G. Robbins, "The CPSC and Textile Flammability," Modern Textile Magazine, April 1976, pp. 7-17.

W. A. Storey, "Energy and Environmental Regulation Impact on the Textile Industry", American Dyestuff Reporter, August 1976, pp. 17-21.

INDUSTRY RESPONSE TO GOVERNMENT REGULATION

David Hillier, American Textile Manufacturers Institute

Abstract

The American textile industry is feeling the deep and serious repercussions of what it calls "The Washington Regulatory Explosion." It is faced with ever-more stringent and expensive environmental and safety and health controls.

In principal, the textile industry does not oppose the goals of these regulations. It, too, promotes a cleaner environment. It works toward a safer and healthier workplace for its employees.

On the other hand, the textile industry believes that it has a responsibility to distinguish between good and bad regulation, both in its own self-interest and the best interests of the nation.

The textile industry believes that every restraint should be most carefully scrutinized and discussed. When the evidence warrants, it opposes bad regulation with all the vigor it commands.

Yet, when a regulation is good, this industry will make every effort to find ways to make it work.

Good afternoon, ladies and gentlemen. My name is David Hillier and I am Assistant General Counsel for Fieldcrest Mills, Incorporated, a diversified manufacturer of textile products, including towels, sheets, blankets, bedspreads, carpets and rugs, with corporate headquarters in Eden, North Carolina.

Fieldcrest is one of many textile firms that is a member of the American Textile Manufacturers Institute. ATMI is the national trade association representing textile companies in 47 of the 50 states. ATMI companies together produce approximately 85 percent of the textile products made in the United States.

It is my pleasure to represent ATMI and the textile industry here today. It is a special honor for me to join the distinguished people on this panel in a discussion of the interdependent interests of consumers, government and industry. My portion of the program focuses on "industry's response to government regulation."

The word "response" reminds me of an exchange between a plumber and one of the federal regulatory agencies in Washington. It seems that the plumber wrote this particular agency saying that he had found hydrochloric acid to be just super for cleaning drains, and, he asked, was it harmless?

The Washington agency replied: "The efficacy of hydrochloric acid is indisputable, but the chlorine residue is incompatible with metallic permanence."

The plumber wrote back saying that he was mighty glad the agency agreed with him. The agency replied: "We cannot assume responsibility for the production of toxic and noxious residues with hydrochloric acid, and suggest that you use an alternate procedure." The plumber was happy to learn that the agency still agreed with him.

Whereupon Washington exploded: "Don't use hydrochloric acid: it eats hell out of the pipes."

I don't know whether this is a true story or not, but it does illustrate a typical response of Washington agencies to problems which exist in our country. The response is illustrative of the way regulatory agencies seem to purposely ignore straightforward and relatively simple answers to complex problems.

The textile industry, as much as any business, has felt the deep and serious repercussions of this regulatory approach. We are faced with ever-more stringent controls on water quality, noise, cotton dust levels, flammability, toxic chemicals and many others.

Many of these controls are, we believe, unnecessary. These controls also present other problems for textiles, and it is my purpose here today to discuss some of these problems with you.

Let me begin by assuring you that the textile industry is not opposed to any and all government regulation per se. This industry's knee does not jerk automatically in violent reaction to all proposed regulation. We believe that "crying wolf" is not an effective way to participate in the political and social process in connection with government regulations.

We do believe, however, that we have a responsibility and an obligation to distinguish between good and bad regulation, both in our own self-interest and the interest of the nation. We also believe we have a duty to make our position known, often done at great expense, when we believe government restraints to be harmful. When a proposed new regulation, however commendable in purpose, would weaken our industry, put thousands out of work, or raise prices beyond people's ability to buy, we believe such a regulation should be most carefully scrutinized and discussed and, if necessary, adamantly opposed with all the vigor we command.

There are several such regulations now facing the textile industry.

The proposed new federal noise standard is an excellent example. The Occupational Safety and Health Administration--OSHA--has proposed a maximum noise level in textile plants of 90 decibels.

It may come as a surprise to some that the textile industry doesn't oppose this goal in principle. We agree that reduced noise levels reaching our workers would provide a more pleasant, safer and healthier working environment.

Where we part company with the regulators is on the means and the cost of achieving this end.

We have taken the position--a very realistic and defensible position--that there is practically no machinery now available which will operate to reduce noise levels in our weaving and spinning rooms to 90 decibels.

We have maintained, and still maintain, that it is presently technologically impossible to attain a 90 decibel standard by means of engineering controls. We have pointed out, even if present equipment were capable of

modification, that the cost would be an estimated \$3 billion. If present textile machinery had to be replaced, rather than modified, the estimated cost in the textile industry for meeting the standard would be \$13 billion.

Even the \$3 billion figure is far in excess of the textile industry's total annual expenditure for new plants and equipment. \$3 billion is also in excess of the industry's total net earnings after taxes. If the government required that this money be spent to convert existing machinery, assuming for the sake of argument this were always possible, hundreds of smaller textile companies would be forced to close. Thousands of people would lose their jobs, and textile growth would come to a standstill.

There is another solution. And it works. We can protect employees now, today, against excessive noise levels. We are already doing it. The method is costly. It involves personnel, time and careful administration. But, it works. It achieves the desired result of reducing textile worker exposure to harmful noise. This solution is a program of personal ear protection combined with a hearing conservation program. This method does not eliminate the noise, but it does prevent the noise from reaching the ears of our people. We are providing personal protective equipment by way of a relatively simple device--ear plugs. More, we are conducting audiometric tests on our employees, continuously measuring their hearing ability.

To us in textiles, this approach seems to be a most reasonable, practical and effective method of meeting the government's objective, pending the day when feasible machinery will be available.

Sometimes I get the feeling that bureaucrats themselves have a hearing deficiency. Our arguments about cost often seem to fall on deaf ears.

Let's take another example. The industry has a problem with cotton dust. It produces a reaction in a few people--by no means all--which causes respiratory difficulties. This condition is called byssinosis. Byssinosis is sometimes called "brown lung" and is unfairly lumped with "black lung," to which coal miners are subject. This is unfair because: there is a correlation between byssinosis and cigarette smoking; because not all cotton textile workers are susceptible to byssinosis; because byssinosis does not progress once a sufferer is removed from the harmful environment; and because the agent in cotton dust which causes byssinosis is not known.

Nevertheless, we want to get rid of byssinosis. So does the government.

OSHA believes the best way to get rid of byssinosis is to reduce the level of respirable cotton dust per cubic meter of air to point-two milligrams. Two-tenths of a milligram of cotton dust would barely fit on the head of a pin. Now imagine this amount of dust dispersed in a whole meter of air--a space larger than a cubic yard.

For the industry to meet the proposed standard of point-two milligrams, the estimated cost is \$2.8 billion--a frightening figure.

The great tragedy of this proposed standard is that it is so unnecessary. We can protect our employees against byssinosis. We are doing it now.

We are measuring the dust level in our plants. We are giving periodic respiratory tests to employees, and we are transferring those few employees who have respiratory problems to areas of our mills where cotton dust is not present. We are providing respirators to some workers. We are installing better filtration equipment.

ATMI has produced a work practices standard which has

been furnished to ATMI member companies, outlining a program of protective measures.

By these means, the industry has proven that textile workers can be effectively and adequately protected from byssinosis.

The textile industry has also taken it upon itself to discover the true prevalence and severity of byssinosis. More than 100,000 tests have been administered to cotton textile workers to determine both their subjective and objective symptoms. The results show that only one percent of textile employees have any such symptoms of the disease. Most of these are simply reactors. . . people who react to cotton dust as some of us react to pollen. Those employees who have any form of disability from byssinosis would be far below that one percent level. Our documented test results--some 2,600 pages--have been submitted to OSHA.

Yet, we still face a multi-billion dollar government imposed dust standard that is unnecessarily stringent.

One more example. The textile industry is required to make children's sleepwear flame resistant. It was the government that ordered the flammability standard, and it was ordered prematurely over the objections of the textile industry. We in textiles repeatedly informed the Congress and the Secretary of Commerce that the necessary technologies and tests for safety and toxicity did not exist. We documented the lack of technologies and cautioned against hasty action without time for thorough toxicity tests. Our pleas again fell on deaf ears.

Just recently, six years after our warnings, the government issued a ban on all children's sleepwear treated with the flame retardant chemical Tris. Tris was the only flame retardant chemical available for polyester at the time the flammability standard was adopted.

Companies in the textile industry did make every test they knew to insure that the flame retardant additives were both effective and safe. And, I repeat, we strongly urged the government to hold off implementing the flammability standard.

Now, because of government action, the industry is faced with an economic loss in the neighborhood of \$200 million. More important, the consumer faces confusion. Mothers are putting their children to bed in their underwear, in hand-me-downs, or in nothing at all.

We in textiles only hope that the consumer knows where the true fault lies and will direct her outrage where it truly belongs.

Each of these regulatory problems I've discussed--and some that I haven't touched upon--have several things in common. First, they are all commendable in their intent. Second, they also involve enormous sums of money, and ignore the damage to this industry of such outlays. Most of them call for technology that doesn't exist.

There is, I think, a third common characteristic of these regulatory problems which has been ignored, overlooked, and misunderstood by both government and most people in our country. That is the ultimate cost to the American consumer.

Suppose, for a moment, that all the proposed regulations affecting textiles were to become law. Imagine, too, that this industry actually spent the billions of dollars required to meet these regulations. Apart from the obvious disaster to the industry itself and its one million employees, what would the impact be on the public?

Obviously, there's no way to precisely measure the probable jump in the price of textiles. No doubt, it would

be astronomical. Just take one example. What happened to the price of children's sleepwear when the flammability rule went into effect? It went up an average of one dollar per garment. Taken by itself, that's not very much. Multiply it by all the children who wear nightwear, and that's still only a drop in the bucket compared to the total cost of this regulation.

Now, try to estimate the ultimate pass-along of the billions of dollars for water preservation, for noise control, for dust reduction, for energy conservation and all the other mandates of government. Who's going to pay?

The obvious answer is the consumer. But is that a correct answer? Suppose all textiles doubled or tripled in price. How many consumers could or would maintain their present buying patterns? How many would have to, or choose to, settle for less, perhaps far less?

No, I don't really expect this to happen, because I believe good, old-fashion common sense will have its say, and we will reach at least some sensible compromise with the government.

This will not happen, however, until the American people become aware of the costs of regulations, and until they recognize the trade-offs involved in cost and a reduced standard of living.

It would be wonderful to have perfectly clear water and pure air. It would be wonderful to have absolutely silent, totally dust-free textile plants which had zero impact on the environment and where no potentially dangerous chemicals were in use, and where everything could be produced without energy. That's idealism at its highest.

But, we must contend with reality. We must participate in the regulatory process without being automatically against all regulation. We must inform the public that compliance with regulations, no matter how beneficial they are, will cost money and the consumer will ultimately pay the tab. Unless we convey this credibly and clearly, the people of this nation will not be in a position to make informed judgements. . .to decide between benefit and cost.

We in the textile industry will continue to cooperate with government in the process of making this nation a better place in which to live and work. We will oppose excessive and unnecessary government regulation. We will point out the effects of bad regulations.

And, when a regulation is good, we will help find ways to make it work.

George B. Sproles, University of Houston
Loren V. Geistfeld, Purdue University

Abstract

This paper presents some major issues and relevant descriptive research findings on consumer satisfaction and dissatisfaction with clothing and textiles. Emphasis is placed on the issues of identifying general levels of consumer satisfaction/dissatisfaction (CS/D), CS/D across specific clothing and textile products, CS/D across general product categories, CS/D with specific product characteristics, identifying the dissatisfied consumer, types of consumer complaints, effects of regulation on CS/D, and consumer satisfaction vs. consumer efficiency. Based on general findings it is estimated that approximately 1 in 5 consumers can indicate dissatisfaction with clothing at any given time, but only a small proportion of those dissatisfied consumers will actually register complaints.

Introduction

A growing amount of recent literature has focused on the problems of defining, conceptualizing and measuring consumer satisfaction/dissatisfaction (CS/D) with products and services. Representative of current approaches to these general issues is a recent collection of papers from a conference conducted by the Marketing Science Institute under sponsorship of the National Science Foundation (Hunt, 1977). The topic has also frequently appeared on programs of the Association for Consumer Research, and was recently the focus of a workshop held at the Indiana University School of Business. These concerted efforts, as well as many individual efforts preceding them, has begun to identify the theoretical, methodological and philosophical issues surrounding investigations of consumer satisfaction and dissatisfaction. The major result is that a relatively well-defined topic area of CS/D is now properly entering the mainstream of contemporary consumer sciences.

This paper will review many of the major issues which can be raised in analyzing CS/D, focusing specifically on the area of CS/D with clothing and textiles. Though many of the general issues regarding CS/D have been identified in recent publications (i.e. Hunt, 1977), it appears that only a limited number of inquiries have focused on the fundamental problems of CS/D within specific product categories such as clothing and textiles.¹ However, it is within such a framework of specific products where the analysis of consumer behavior becomes most meaningful, especially to the pragmatically-focused interests of marketers, public policy makers, consumer affairs professionals and not least of all the ultimate consumer.

There are two specific objectives of this paper. First, we will identify key issue areas related to analyzing CS/D with clothing and textile products. Each identified issue is selected as a key issue because it is a current area of interest among policy makers, business-

men, consumer advocates or researchers. Second, for each issue area we will summarize some recent research and substantive general findings which answer questions raised by the issue. While our emphasis is on clothing and textiles, many of the issues and points discussed are general in nature and can therefore be very relevant to the analysis of consumer satisfaction and dissatisfaction with other categories of consumer products and services.

An Overview of Key Issues

Many issues and sub-issues can be identified. In this paper we will focus on these key areas: 1) What is the magnitude of consumers' general satisfaction and dissatisfaction with the overall product category of clothing and textiles? Sub-issues under this general issue would include CS/D with general prices/price trends, levels of quality, styles/assortments available, types of stores retailing the products, sizes available and so on. Another major sub-issue can focus on identifying the major component parts of the clothing-textile system which are most likely to cause consumer satisfaction or dissatisfaction (i.e. performance of fibers, fabrics, construction, seams, buttons, zippers or other major component parts). 2) To what extent are consumers satisfied with specific products (i.e. jeans, shirts, suits)? Are certain classifications of merchandise more subject to CS/D than others (i.e. categories of men's, women's and children's wear)? 3) How satisfied/dissatisfied are consumers with clothing and textiles as compared to other major product and service categories? Is dissatisfaction relatively more or less serious than in other product categories? 4) What product characteristics of the clothing - textile system in its actual end use are most likely to result in consumer satisfaction or dissatisfaction over the "wear life" of the product? For example, will CS/D result from changes of style-fashionability, economy, ease of care, durability, comfort, dimensional stability or other expectations regarding performance of the product? 5) Who are the dissatisfied consumers? Do they have a unique profile in terms of general orientations toward consuming clothing, lifestyle, and demographic characteristics? 6) What types of clothing and textile complaints are actually registered with retailers and/or manufacturers? How satisfactorily are these handled? What are the principal reasons for those complaints? To what extent does the complaining behavior of consumers represent actual levels of satisfaction/dissatisfaction in the market? 7) How do governmental regulations of clothing and textiles affect CS/D? To what extent is federal regulation more effective than voluntary self-regulation by industry in providing ultimately for consumer satisfaction? 8) Should the clothing and textile market (or any sub-market) be considered to be performing efficiently (or in the consumer's or society's interest) if consumers are satisfied? Is CS/D a legitimate criterion for assessing market performance? Or can a market be performing poorly even when consumers are satisfied? What are the implications for policy makers?

Each of these issues will be briefly discussed in turn. Our treatment of these issues is not meant to be exhaustive, but rather is designed to introduce some major questions and current thought in each area.

¹Several papers in Hunt (1977) focused on housing and foods, and several basic issues relative to those products were conceptualized. See papers by Hempel, Morris, and Handy in that volume.

General Indicators of CS/D

A variety of recent investigations have provided broadly general indications on the magnitude of consumers' satisfaction/dissatisfaction with clothing and textiles. Research conducted by Sproles under sponsorship of the Indiana Cooperative Extension Service has examined a variety of clothing orientations among adult women living in Indiana (Sproles, 1977). The purpose of the investigation was to survey women's interests/priorities in consuming clothing, information seeking behaviors, and purchasing behavior. A number of aspects of this study focused specifically on issues related to CS/D.

The study involved administering a mailed eight page questionnaire to a stratified random sample of 2,000 households in Indiana. The questionnaires were mailed to the adult female head of each household, and two

follow-up mailings including a postcard reminder and a second copy of the questionnaire were used to increase response rates. A total of 989 complete and usable questionnaires was returned. This sample was found to be demographically representative of households in Indiana (and in the U.S. as a whole), with the exception that older and lower income groups were not fully represented. A more complete discussion of the methodology can be found elsewhere (Sproles, 1977).

Our principal interest is with those findings dealing with CS/D. Of the 150 measures included in the survey, approximately 30 dealt either directly or inferentially with general issues of CS/D. Table 1 includes findings for selected Likert-scaled items measuring consumer interests and priorities regarding clothing in general. These items appear to be relevant to satisfaction/dissatisfaction in particular.

TABLE 1

SOME GENERAL INDICATIONS OF CONSUMER SATISFACTION/DISSATISFACTION WITH CLOTHING

INDICATORS	Percent Responding					
	No Response	Strongly Disagree	Disagree	In-Between	Agree	Strongly Agree
Indicators of Satisfaction:						
<u>I usually have one or more outfits of the very latest style.</u>	.2	5.1	19.9	29.4	36.4*	9.0%*
<u>I keep my wardrobe up-to-date with the current fashion.</u>	.6	6.7	19.8	51.0	19.0	2.9%
<u>I am very well informed about current clothing and fashion trends.</u>	0	5.4	17.5	40.8	26.8	9.5%
<u>To me, shopping for clothes is a pleasure.</u>	.1	4.3	15.4	30.4	34.4	15.4%
Indicators of Dissatisfaction:						
<u>Keeping up with changing fashions is too expensive.</u>	.1	2.5	7.7	19.5	37.7	32.5%
<u>I carefully watch how much I spend on clothes.</u>	.7	1.4	3.1	16.1	41.6	37.1%
<u>In general, the quality of clothing in the stores is lower than in the past.</u>	1.0	6.9	15.3	17.8	29.4	29.6%
<u>People are too concerned about their dress.</u>	1.0	6.7	24.7	32.7	23.8	11.2%
<u>Buying clothes out of catalogs without actually seeing them, is too risky.</u>	.4	4.9	14.9	19.6	29.8	30.4%
Indicators of Priorities for CS/D:						
<u>Clothing style is more important than price.</u>	.7	12.1	26.0	30.8	20.8	9.5%
<u>Clothing comfort is more important than style.</u>	.7	.6	3.8	25.6	47.1	22.1%
<u>Clothing quality is more important than price.</u>	.8	1.7	6.8	30.3	42.2	18.2%
<u>Clothing quality is more important than style.</u>	.6	1.9	9.5	38.3	37.6	12.0%

*Read: Of a sample of 989 women in Indiana, 9.0% "strongly agreed" with the statement at the left, 36.4% "agreed" with the statement at the left, etc.

The first three items in Table 1 give an indication of consumer satisfaction with current fashions. These items would appear to imply that perhaps less than 10 percent of consumers would qualify as being highly satisfied with current fashions, though it would also appear that one-third or more of consumers could be considered at least more satisfied than dissatisfied.

The last indicator of satisfaction in Table 1 relates to satisfaction with shopping for clothing. This item indicates that consumers are for the most part pleased with their experiences in shopping for clothes.

The next five items of Table 1 can be classed as indicators of consumer dissatisfaction with clothing and textiles. Collectively these five items imply a substantial degree of general dissatisfaction with clothing in one-third or more of the adult female population. The respondents clearly see changing fashions as expensive, and they view the quality of clothing as lower than in the past. Responses to the item "People are too concerned about their dress" may indicate a dissatisfaction with the social pressures toward conformity in dress. Finally, the last item suggests that many consumers would perceive substantial risks of dissatisfaction with clothing purchased "blind" from catalogs.

The last four items in Table 1 can be termed indicators of priorities for consumer satisfaction/dissatisfaction. Items similar to these can be used to interpret general trends in the determinants of consumer satisfaction. For example, the four items appear to suggest that many consumers would place priority on clothing comfort and "quality" over style and price in judging their satisfaction with clothing and textiles.

Table 2 summarizes problems with clothing reported by consumers participating in the study. The table clearly indicates a great proportion of consumers have experienced dissatisfaction with major physical features or performance characteristics of garments. The most significant area of problems appears to be in construction, with 68 percent of the respondents reporting recent experiences with seams tearing apart. Also notable is the fact that 52 percent reported "lower overall quality" to be a problem with recently purchased clothing.

Table 3 presents consumers' ratings of the importance of selected criteria in influencing their purchases of women's outerwear. The overall findings of this table are that many styling, performance and price features are often or always important to consumers. Only marketing and shopping factors (i.e. brand name, salesperson's advice) and the "most current fashion" were rated important by a small proportion of consumers, and even these factors were clearly important to as much as one-quarter of the subjects. Probably the most important contribution of these findings is to show that most consumers' expectations regarding their clothing and textile purchases are extremely high. They expect their purchases to be almost "perfect" over a wide range of characteristics, and it is not hard to imagine that consumer dissatisfaction could result with frequency under such high expectations.

Several other recent studies provide general indications of CS/D. A recent paper by Plummer (1977) reports findings of a series of studies using items similar or identical to those in Table 1. Different surveys using these items were conducted longitudinally from 1967 to 1975. Table 4 summarizes the percentage of subjects agreeing with each statement for each survey. Each of the three items indicates a clearly declining trend of agreement over the years. Plummer interprets these items to indicate that shopping is becoming more of a "hassle" and that consumers are becoming more casual in their attitudes/approach to dress. He also suggests that measurement of such trends should be a part of a

TABLE 2
PROBLEMS WITH CLOTHING PURCHASED
IN THE PAST YEAR

	PERCENT OF CONSUMERS REPORTING PROBLEM
Poor fit	48%
Fabric wearing out too quickly	40
Colors fade in washing	33
Cleaning instructions not easy to understand	11
Seams tear apart	68
Mistakes in construction	57
Shrinkage when washed	44
Zippers not working	40
Lower overall quality	52

Note: Respondents were allowed to check as many problems as they had experienced.

TABLE 3
IMPORTANCE OF SELECTED CRITERIA IN
CLOTHING PURCHASING DECISIONS:
WOMEN'S OUTERWEAR

CRITERIA FOR CLOTHING PURCHASING DECISIONS	PERCEIVED IMPORTANCE OF SELECTED CRITERIA		
	Seldom or Never Important	Sometimes Important	Often or Always Important
<u>Styling</u>			
Style looks good on my figure	0.5*	2.4*	96.7*
Color of the garment	0.9	10.8	87.9
Pattern or design of the fabric	2.8	11.8	84.7
Conservative styling	8.3	27.4	63.8
The most current fashion	30.7	43.2	25.6
<u>Product Features</u>			
Comfort of the garment	.4	2.0	97.0
Ease of care (washing, cleaning)	1.1	4.4	93.9
Quality of construction	1.5	4.4	93.2
Cost of the garment	1.2	10.2	88.3
Garment will wear for a long time	2.3	17.4	79.8
Fiber content of the fabric	9.2	20.6	69.7
<u>Market/Shopping Factors</u>			
Reputation of the store	7.4	24.3	67.6
Opinion of my shopping companion	23.2	46.0	30.0
Brand name of the garment	32.0	39.2	28.3
Advice from the store's salesperson	34.3	39.2	25.9

*Read: Of a random sample of 989 women throughout Indiana, 0.5% indicated that "style looks good on my figure" is seldom or never important to their clothing purchasing decisions; 2.4% indicated this criterion to be sometimes important; and 96.7% indicated this criterion was often or always important. Rows do not add to 100% due to missing answers by a small number of respondents.

TABLE 4

CS/D Across Specific
Clothing and Textile Products

STATEMENT	TRENDS IN SELECTED CLOTHING ORIENTATIONS				
	PERCENT AGREEING, BY YEARS				
	1967	1968	1971	1973	1975
I usually have one or more outfits of the very latest style	30	31	30	23	20%
When I must choose between the two, I usually dress for fashion, not for comfort	18	19	15	13	12%
I love to shop for clothes	58	56*	52	46	47%

*1969 Survey

Source: Plummer (1977, pp. 395, 397)

broad framework for analyzing CS/D from a variety of perspectives.

A series of consumer surveys conducted in Syracuse, N.Y., have examined a variety of dimensions of CS/D with both clothing and home furnishings. In the first study involving a sample of 193 returns from a mail survey, Steiniger and Dardis (1971) found most respondents were satisfied with the performance of clothing and textiles. Over 70 percent of subjects were above average in satisfaction, and only 10 percent were below. However, although considerable satisfaction was noted, respondents still reported many problems with durability, appearance, ease of care, and comfort. Sixty-seven percent of the problems with both clothing and home furnishings were associated with durability. These were predominantly in construction of fabric failure (i.e. pilling, seam tears, worn areas), and to a much lesser extent problems with fabric stretching, shrinking, or color change. Finally, a substantial number of consumers mentioned poor performance of "no-iron" clothing and difficulty in cleaning home furnishings.

The second Syracuse study focused on home furnishings made from textiles, and involved a sample of 908 responses to a mailed questionnaire. In this study Nichols and Dardis (1973) found 66 percent of the respondents to be relatively satisfied with their purchases of home furnishings during the past years. However, as in the previous study there were a substantial number of problems reported among dissatisfied consumers who were identified. The majority of those problems (53 percent) involved appearance and ease of care (i.e. soiling, difficulty in cleaning, color changes), while 24 percent were problems of durability. Finally, it is especially notable that some dissatisfaction was expressed for one-fourth of all items purchased.

A follow-up survey was conducted on the 280 dissatisfied consumers in the preceding study. In this follow-up 50 percent of the subjects indicated dissatisfaction with wear and durability, and 44 percent indicated dissatisfaction with appearance and ease of care. This divergence from the earlier findings was explained by the fact that the follow-up focused only on those consumers who reported dissatisfaction, while the previous findings included "satisfied" consumers who still reported "problems". Presumably the problems reported by satisfied consumers were not extremely serious. It would therefore appear that the main concern of the dissatisfied consumer is with durability.

The preceding studies give some broad indications of consumer satisfaction and dissatisfaction with the generic category of clothing and textile products. A number of studies have also focused on analyzing CS/D with specific classifications of merchandise and specific items of clothing. A variety of research approaches ranging from surveys to actual consumer wear tests have been conducted, and the following paragraphs will summarize some approaches and findings of these investigations.

A series of major surveys of consumers' perceptions of "value for the money" received in today's market has been conducted by National Family Opinion, Inc., under sponsorship of The Conference Board (Linden, 1974). The 1974 survey included questions on 45 different products and services, and was administered to a representative sample panel of 10,000 households. Panel members were asked to report if value for the money was "good," "average" or "poor" for the various products and services. Table 5 summarizes findings for clothing and textiles. As Linden (1974) points out, these data show clear dissatisfaction with the value received in clothing and textiles. Of special interest presently is the wide variation in dissatisfaction from one general classification to the next. Men's clothing appears to offer the most satisfaction, but both women's and children's clothing are rated a "poor" value by about half the sample. Regarding home furnishings, satisfaction with carpets is highest for all textile products, but satisfaction with upholstered furniture is comparatively low. Finally, Linden notes a decline in consumer satisfaction with each category as compared to a similar survey conducted a year earlier (Linden, 1974).

Another perspective of CS/D with specific lines of clothing and textiles is offered in research by Hughes (1977). He focused on analyzing consumers' satisfaction with specific purchases reported by a sample of 928 consumers in a large midwestern city. Table 6 summarizes some key findings. In general, about 80 percent of consumers were completely satisfied with their purchases. However, satisfaction clearly varied by lines of merchandise. For example, about 95 percent of the purchasers reported complete satisfaction with men's dress slacks, while only 74 percent were completely satisfied with women's jeans. But perhaps the most significant findings are in the relation between price paid and consumer satisfaction. For 16 of the 22 purchase classes, people who had lower than average satisfaction with their purchases reported a distinctly lower average price paid than those who were completely satisfied.

TABLE 5

CONSUMERS' PERCEPTIONS OF VALUE RECEIVED FOR THE
MONEY IN SELECTED CLOTHING AND TEXTILE PRODUCTS, 1974

CATEGORY	PERCENT RESPONDING		
	Good	Average	Poor
Men's Suits	10.0	61.7	28.3%
Women's Dresses	6.8	48.2	45.1%
Children's Clothing	6.2	43.6	50.2%
Shoes	8.1	45.7	46.1%
Upholstered Furniture	7.2	56.5	36.3%
Carpets	15.9	67.2	16.9%

Source: Linden (1974, p. 52)

TABLE 6

CONSUMER SATISFACTION WITH CLOTHING AND TEXTILES,
AND PRICES PAID BY SATISFACTION LEVELS

PRODUCT	Number of Purchases	Percent Completely Satisfied with Product	Average Price Paid	
			Completely Satisfied Consumers	Consumers Satisfied Lower than Average
Men's Dress Slacks	136	94.9%	\$ 17.35	\$ 17.86
Men's Knit Sport Shirts	157	86.6	11.59	7.41
Men's Reg. Sport Shirts	148	85.1	7.58	8.17
Men's Casual Slacks	255	80.4	13.95	11.40
Men's Dress Shirts	382	84.3	9.07	8.79
Men's Suits	257	82.5	109.81	106.36
Men's Shoes	377	80.6	29.00	22.25
Women's Dresses	429	82.3	36.56	26.17
Women's Bras	464	81.5	5.57	4.10
Women's Tops	494	81.0	11.28	9.31
Women's Pants/Jeans	511	74.0	15.79	11.44
Women's Robes	173	83.2	14.59	15.35
Women's Panty Hose	499	71.9	2.23	1.76
Boys' Shirts	194	79.5	6.04	4.68
Boys' Slacks/Jeans	232	82.3	7.94	7.24
Girls' Pants/Jeans	83	81.1	10.15	8.98
Girls' Dresses	156	79.9	21.07	13.09
Carpeting	139	74.8	722.29	549.72
Sheets	335	83.3	5.93	5.83
Draperies	221	80.1	162.82	93.77
Curtains	219	79.0	27.23	28.57
Bedspreads	222	78.8	25.07	21.32

Source: Hughes (1977, pp. 314,315,331)

One of the most direct ways of measuring consumer satisfaction with specific products is by conducting actual consumer wear tests. A wear test involves use of a panel of consumers who are given selected brands of a product to actually wear for a period of time, with evaluations of performance made at different stages of the wearings. Such studies are frequently conducted by home economists and industrial producers of products. For example, a study by Stover (1972) focused on consumer satisfaction with durable press shirts, and found that consumers were most satisfied with lighter weight polyester/cotton blends in shirt fabrics. Similarly, a study of consumer acceptance of flame retardant sleepwear, Laughlin and Buddin (1977) found consumers preferred a 65/35 percent blend of cotton/polyester over other tested fabrics.

Such wear test can ultimately have the most "face validity," since consumer satisfaction is actually based on measuring consumers' actual use of the product under controlled experimental conditions. This is one reason why such studies are frequently conducted. However, wear studies have some limitations. In most cases it can be either prohibitively expensive or time-consuming to wear test a complete array of competitive products with a large enough sample of consumers to be statistically representative of the market. Further, a complete wear test may take a long period of time, since the subject must actually wear the garment in his or her "normal" manner. In spite of these obvious limitations, wear tests are thought to be beneficial in forecasting consumer satisfaction, and will continue to be used as a measure of satisfaction.

CS/D Across Product Categories

The next issue which naturally arises focuses on the degree to which consumers are satisfied/dissatisfied with clothing and textiles in comparison to other major categories of consumer products and services. One viewpoint

holds that clothing and textiles is not a principal area of concern in comparison to other areas of consumer problems. For instance, Margolius (1975) has indicated that major problems faced by consumers include higher costs of food, medical care, housing, loans, repairs, and autos, but clothing and textiles are not mentioned. Furthermore, the recent Sentry Insurance study (1977) of consumer issues surveyed numerous consumer concerns and found considerable general concern with prices, qualities, and consumer information, but clothing and textiles was never specifically identified as a problem area. These examples might imply that in the arena of CS/D, there is bigger game than clothing/textile products.

It is quite plausible that many areas of consumer problems deserve higher priority than consumer dissatisfaction with clothing and textiles. Nevertheless, there is some evidence that consumer dissatisfaction with clothing and textiles is more substantial than for many other categories of products and services. For instance, in the National Family Opinion study introduced earlier, consumer satisfaction with women's dresses, children's clothing shoes and upholstered furniture was substantially lower than average satisfaction for all items (Linden, 1974). Most items of foods (meats and fresh vegetables), appliances, transportation, life insurance, and TV rated substantially higher than clothing and textiles in general, and only repairs consistently rated lower. Only the categories of men's suits and carpets rated above the average for all items.

A few other studies give some general indications of CS/D across product categories. In one study reported by Plummer (1977), consumers offered judgments on whether or not product quality had improved in the past 20 years. A total of 50 percent of respondents reported clothing as "not as well made." In comparison, products receiving a larger proportion of "not as well made" ratings were houses (77 percent) and autos (70 percent), and those

receiving a lower proportion were household appliances (34 percent), small appliances (31 percent), and TV sets (20 percent). In a 1975 study cited by Plummer (1977, p. 387), clothing was rated as an excellent or good value for the money by 27 percent of consumers. In comparison, lower rated products were gasoline (12 percent), prescription drugs (21 percent), and automobiles (22); higher value for the money ratings went to electrical appliances (37 percent) and fresh foods (39 percent). Data from the Hughes (1977) study also indicate that consumers are generally more satisfied with certain major purchases (i.e. TV, washer, dryer), but for other major purchases the proportion of satisfied purchases generally equals averages for clothing (i.e. furniture, stoves, mattresses, refrigerators, tires, batteries). Finally, in a study by Pfaff (1976), clothing, food and "standard of living" were virtually tied in consumers' ratings of overall satisfaction, but satisfaction with appliances was significantly higher. The conclusion which follows from these studies is that clothing satisfaction among consumers suffers in comparison to a variety of other major purchases, though it is far from being the "worst" area of consumer dissatisfaction.

CS/D With Product Characteristics

Little recent published research has focused on consumer satisfaction as a function of specific product characteristics (attributes) such as styling, durability, ease of care and so on. However, some general principles have become established. Previous investigations by home economists have suggested that for clothing to be satisfactory it must rate high on the product characteristic (s) most important to the consumer, and it also must meet certain "minimum" expectations on other characteristics (Ryan, 1966, pp. 147-152, 178-187). In other words, the consumer must in most cases be relatively satisfied across a number of product characteristics in order to receive overall satisfaction. See Table 3 for current evidence on this point.

Another general principle has been reinforced in a recent exploratory investigation by Swan and Combs (1976). They focused on the contrasting roles of instrumental and expressive product characteristics in influencing CS/D with clothing. Instrumental characteristics were defined as those involving physical performance of the product (i.e. durability, ease of care), while expressive characteristics were related to "psychological" levels of performance (i.e. styling, fashionability). It was found that dissatisfaction with clothing was likely to occur when instrumental characteristics did not meet the consumer's expectations, while satisfaction was most associated with fulfilled expressive expectations. The findings also suggested that the consumer's instrumental requirements must be satisfied before expressive requirements if overall satisfaction is to occur. The major implication of these findings is that:

In judging the performance of a product, the consumer compares a set of performance outcomes to the outcomes that were expected for the item. If the performance of the physical product was below expectations, then the product is likely to be categorized as dissatisfactory. If both instrumental and expressive outcomes were equal to or exceeded expectations, then the consumer will tend to judge the product as satisfactory. (Swan and Combs, 1976, p. 33)

Findings cited earlier in this paper, which showed consumer dissatisfaction to focus mainly on physical performance failures (i.e. failures in construction, durability, ease of care) would seem to lend further support to this proposition.

Identifying the Dissatisfied Consumer

The varied studies we have reviewed indicate a substantial proportion of consumers may be dissatisfied with different aspects of their clothing and textile purchases. Who are these consumers, and what are the characteristics of consumers who are most likely to voice their complaints? There are several studies of consumer dissatisfaction which indicate the dissatisfied consumer has some important characteristics. For example, studies by Liefeld, Edgecombe and White (1975) and Warland, Hermann and Willits (1975) suggest that the consumer who voices dissatisfaction with a purchase is well educated, young, relatively high in income, and above average in social class and group membership. Furthermore, it was found that there are some characteristics of those who experience dissatisfaction but do not voice it. These consumers, according to Warland, Hermann and Willits (1975), are lower income, less well educated individuals who are not involved in consumer and political activities. This finding is supported by Best and Andreassen (1976), who found that households in lower socioeconomic classes tend to perceive fewer problems with their purchases of goods and services and voice their dissatisfaction less frequently than do consumers in higher socioeconomic classes.

However, a number of studies have found that dissatisfied clothing and textile consumers have no distinct demographic or socioeconomic characteristics (Linden, 1974; Nichols and Dardis, 1973; Steiniger and Dardis, 1971). Though some of these findings modestly imply that dissatisfied consumers are in higher income and educational groups, it appears for the most part that dissatisfied clothing consumers are found in all ages and levels of the consumer population.

A preliminary analysis of data from the study of clothing orientations among adult women in Indiana conducted by Sproles (1977) gives more evidence on characteristics of dissatisfied consumers. To conduct the analysis, a general index of consumer dissatisfaction with clothing was formed for each subject by summing four indicators of dissatisfaction presented earlier in Table 1 (the item "I carefully watch how much I spend. . ." was not included in the index). The index was divided into subjects scoring "high" (top one-third of scorers) and those scoring "medium to low" (remainder of subjects). These two groups were cross-tabulated against 27 clothing interests/priorities measures, 24 lifestyle activities, 15 purchasing criteria, consumption indexes for sewing and purchasing, fashion interest, and 11 demographic and socioeconomic measures.

Findings of this preliminary analysis tentatively indicate that the dissatisfied consumer is no different from other consumers in lifestyle activities, demographics and socioeconomic characteristics. There were some tentative indications that the dissatisfied consumer might be concentrated toward younger, middle income groups, but this finding was not significant. It was also found that dissatisfied consumers are no different from other consumers in their level of consumption (purchasing and sewing of clothing). However, it was found that dissatisfied consumers differed significantly from others on a variety of clothing interests and priorities. Specifically, dissatisfied consumers were more likely than others to agree with a variety of statements indicating they are generally price and quality conscious. There were also indications that dissatisfied consumers may be lower in fashion interest than others. Finally, they were significantly more likely than others to rate the specific purchasing criteria of garment cost, choosing garments that wear for a long time, and comfort of the garment as "always important." In short, there appears to be a substantial orientation toward economy-consciousness and quality-consciousness among those consumers who are most likely to report dissatisfaction with clothing.

Types of Consumer Complaints

It is an accepted principle that although many consumers may be dissatisfied with the performance of a product, few will actually submit a formal complaint seeking redress to either the retailer or manufacturer. Nevertheless, those complaints which are actually voiced are frequently thought to be at least modestly representative of the problems consumers actually experience with products and services. Thus actual voiced complaints can be viewed as potential indicators of market performance, areas needing the attention of producers, and problems requiring governmental intervention (regulation) of markets.

It is first useful to summarize some of the major areas in which consumers will voice complaints about products and services. In a broad study of all types of product/service complaints made by consumers, Diamond, Ward and Faber (1976) found that 14 percent encountered a pre-purchase problem, 27 percent experienced a purchase/transaction/delivery problem, 17 percent were unhappy with product performance, 17 percent encountered a guarantee/warranty/contract problem, 26 percent were unhappy with the service/repair rendered and 4 percent encountered a problem with deposits/credit/collections. Of these complaints 34 percent involved specific retail outlets with about 9 percent, each, of these complaints directed towards department stores, clothing stores and discount stores.

In what is perhaps the most comprehensive recent study of consumer dissatisfaction with market performance and subsequent complaining behavior, Best and Andreasen (1976) found that 83.4 percent of the respondents were completely or partially satisfied with their clothing purchases while 13.9 percent experienced some degree of dissatisfaction. When the respondents were asked what types of problems they experienced with clothing purchases, slightly over 6 percent indicated price-related problems while about 28 percent indicated non-price problems (principally problems with product performance). Best and Andreasen note further that across all product and service categories analyzed, 3.1 percent of the perceived problems related to shrinking or fading, 23.2 percent to workmanship, 7.1 percent to fit or size and 6.2 percent to stitching. With respect to clothing, 7.7 percent of all purchases resulted in problems related to poor workmanship, 9.7 percent to price and 30.4 percent to other types of problems. Of those clothing consumers who experienced problems with a purchase, 65.8 percent took no action, 7.0 percent changed shopping patterns, 26.0 percent voiced their problem to the seller and 1.0 percent voiced their problem elsewhere. For those consumers who voiced their dissatisfaction with clothing purchases, 75.3 percent achieved satisfactory redress of their problem while 18.7 percent considered redress received as unsatisfactory. (Best and Andreasen, 1976, pp. 16, 27, 42, 89)

Other investigations have indicated that although consumers may experience dissatisfaction with purchases, only a small proportion will actually register complaints. For example, Steiniger and Dardis (1971) found 56 percent of respondents to a mail survey cited specific performance problems with clothing and linens, but less than one in four of these consumers (23 percent) actually registered complaints. Complaints which were registered centered on problems with durability (73 percent of complaints, on specific problems of construction, stretching/shrinkage, color changes), with appearance and ease of care being the other major problem area (23 percent of complaints). Similarly, of the 34 percent of respondents reporting problems with home furnishings, less than one-fourth (22 percent) registered complaints. For this category complaints also centered on durability (79 percent), with appearance and ease of care being the other major problem area (15 percent).

The more recent study of consumer satisfaction with home furnishings by Nichols and Dardis (1973) further confirms the preceding findings. Just over one-fourth of the dissatisfied consumers registered complaints, with problems relating to durability, appearance and ease of care being predominant. More than half of the complainants were not satisfied with the store's action taken, and most of those who did not complain indicated they felt nothing would be done or it was too much trouble to complain. Finally, they found that the higher the price paid for an unsatisfactory item, the more likely the consumer will register a complaint. This very important finding is consistent with general findings of other researchers (Best and Andreasen, 1976).

Effects of Regulation on CS/D

In the past consumer satisfaction/dissatisfaction with clothing and textiles was probably based primarily on the degree to which products met the consumer's expectations on characteristics including fashionability, quality and economy. However, with the recent growth of market regulation such areas as product safety, product labeling and product warranties have been legitimately installed as criteria for CS/D. However, for these and many other regulatory issues there is currently little empirical evidence indicating the impact of regulations on CS/D with clothing. Nonetheless, we may hypothesize as to how CS/D is influenced.

Perhaps the most controversial issue in regulation involves the safety of clothing products. Recent concerns with fabric flammability and carcinogenic chemical finishes (i.e. the TRIS case) have been dominant in the clothing trades and have received a great deal of publicity. In the past the safety of such items as "platform" shoes and loosely fitted clothing has been questioned both by consumer advocates and medical authorities. At the very least the publicity received by these and other topics has increased consumers' awareness of new clothing problems which involve personal health. This may lead to a variety of results, i.e. fear of the product, fear of specific brands, boycotting of products, and even a general suspicion of clothing products. The ultimate result can be increases in consumer dissatisfaction with the current market.

Product labeling is another current issue among consumer interests. In this area the prospects are that the growing emphasis on informative labeling, both by government and industry, will lead to greater consumer satisfaction. The recent Sentry Insurance study (1977) found that about 18 of every 25 consumers believe there has been an improvement in informative labeling of products. From another perspective, in a survey of consumer attitudes toward nutritional (content) labeling Lenahan et. al. (1972) found that 96.5 percent of the respondents felt nutritional labeling was a good idea, 50.9 percent said they use or would use the labels while 9.2 percent actually used them; however, about 85 percent of the respondents indicated there were positive nonuse benefits, i.e. nutritional labeling would enhance consumer confidence in the food industry, thus increasing satisfaction. Extrapolating these findings to informative labeling for clothing, it is likely consumers would perceive a clothing market with readily accessible informative labeling as presenting a less hostile environment in which to operate regardless of whether or not they actually used the informative labels. Therefore, one could expect informative labeling to reduce the incidence of consumer dissatisfaction, and increase satisfaction.

In the area of informative product labeling, care labeling of clothing and textiles has recently been the central area of concern. Research by Steiniger and Dardis (1971) has indicated informative care labels and tags on clothing can increase consumer satisfaction, probably

because those using the information can minimize damage done to the product in care and cleaning. However, there has recently been extensive discussion as to whether regulations governing permanent care labels attached to clothing and textiles have been a success or failure (Powderly, 1976). Generally the information on care labels has been found to be clear but incomplete, and more detailed instructions have been called for. There has also been a problem with "low labeling," which can mislead a consumer by suggesting one care method when another easier or less costly method might work just as satisfactorily. As a result of these types of problems with care labels, the Federal Trade Commission is currently recommending some major changes in labeling regulations.

Warranties are a third area of general concern, especially since the Magnuson-Moss Warranty Act of 1975 was passed. Only a very few major textile fiber and clothing manufacturers offer what is overtly named a "warranty," but many retailers make it a promoted policy to accept returns of unsatisfactory merchandise. However, there is some question as to whether such warranties can increase satisfaction, and in some cases their effect may be to decrease satisfaction. For instance, the Sentry Insurance study (1977) found over one-half of all consumers surveyed felt warranties were inadequate, nearly fifty percent felt warranties were written mainly for the protection of manufacturers, and more than one-half felt warranties were not understandable. Further, as Gerner and Bryant (1977) have noted, for those manufacturers who have below average warranties, there may be an increase in dissatisfaction. Thus it might seem that if providing consumer satisfaction is a major reason for offering a warranty, some manufacturers or retailers might actually not offer warranties in order to avoid the risks of causing consumer dissatisfaction. It can also be noted that the legal restrictions on warranties, which have been viewed as confusing or complex by many businessmen, might only further encourage that trend. Certainly such results would diminish the force of warranties as an ultimate guarantee of consumer satisfaction.

As indicated earlier, the preceding observations must have the flavor of speculation, for only a limited amount of sophisticated research on the effects of these and other regulatory policies has been done. It should also be pointed out that many actions regarding safety, informative labeling and warranties are voluntarily taken by manufacturers either as a competitive marketing strategy or for other related reasons. Thus more attention to these issues is required, to measure the effectiveness of various voluntary business policies vs. mandatory regulations in providing consumer satisfaction.

Consumer Satisfaction Vs. Consumer Efficiency

The last issue we will briefly introduce regards the extent to which consumer satisfaction implies that a consumer is performing efficiently in choosing products available in the market. As we have noted elsewhere (Sproles, Geistfeld and Badenhop, 1977) the concept of consumer efficiency implies that consumers obtain the maximum possible satisfaction (utility) given available resources. This concept sets a norm against which observed consumption (satisfaction) can be evaluated. Underlying this notion of consumption efficiency is an implicit assumption that consumers operate under full information. Once this is recognized, it becomes apparent that consumer indication of satisfaction with experiences in the marketplace is only a necessary but not a sufficient condition for the existence of consumption efficiency. If perceived satisfaction were based on ignorance, it is possible that upon removal of the ignorance the consumer who once indicated satisfaction may no longer be satisfied. The policy implications of this

are quite clear. If a consumer indicates dissatisfaction consumption efficiency has not been achieved and corrective action may be warranted. If a consumer indicates satisfaction, the extent to which this may or may not indicate consumption efficiency depends upon how stable perceived satisfaction is upon providing additional information. Should perceived satisfaction prove to be unstable, corrective action may also be warranted.

Conclusion

A number of descriptive studies indicate that consumer dissatisfaction with clothing and textiles is widespread, and that this product category may cause greater consumer dissatisfaction than many other major product categories. As a general and perhaps conservative estimate, it would appear that at any given time 1 in 5 consumers (20 percent) is aware of and can report a specific dissatisfaction, usually involving a physical performance characteristic. However, only a small proportion of the dissatisfied consumers, probably less than 1 in 4, may actually register a complaint with the retailer or manufacturer. Thus complaints may not reflect the full magnitude of consumer dissatisfaction, though they may be a proxy measure of the types of problems consumers frequently find dissatisfying.

It would appear that improvement of consumer satisfaction might be made through several strategies. Consumer educational and informative labeling programs could be used to disseminate appropriate information needed by consumers to make satisfactory decisions. In some cases regulation might be warranted to insure all manufacturers and retailers would offer products fulfilling selected critical criteria (i.e. safety, information on product care/use). Finally, programs of voluntary and/or mandatory product standards which are receiving increased discussion in clothing and textile trades (i.e. see ASTM Standardization News, December 1976) may be the long-run solution to problems of product performance and associated dissatisfactions which consumers now report.

There have been a number of exploratory and descriptive research investigations on the issues we have raised. However, more quantitative findings are needed on a number of the issues before well-founded policies can be implemented by either government and business. But researchers should also be warned of the conceptual and methodological difficulties involved in measuring the phenomenon of CS/D (Hunt, 1977).

References

- Arthur Best and Alan R. Andreasen, "Talking Back to Business: Voiced and Unvoiced Consumer Complaints," Working Paper, Center for the Study of Responsive Law, Washington D.C., 1976.
- Steven L. Diamond, Scott Ward and Ronald Faber, "Consumerism: Analysis of Calls to a Consumer Hot Line," Journal of Marketing, 40 (January 1976), 58-62.
- Jennifer L. Gerner and W. Keith Bryant, "Consumer Economics of Appliance Warranties," Working Paper, Department of Consumer Economics and Public Policy, Cornell University, 1977.
- Donald A. Hughes, "An Investigation of the Relation of Selected Factors to Consumer Satisfaction," in H. Keith Hunt, ed., Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction (Cambridge, Massachusetts: Marketing Science Institute, 1977).

- H. Keith Hunt, ed., Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction (Cambridge, Massachusetts: Marketing Sciences Institute, 1977).
- Kenneth C. Laughlin and Frances Buddin, "Consumer Acceptance of FR Sleepwear: Part II," Textile Chemist and Colorist, 9 (September, 1977), 70-74.
- R. J. Lenahan, J. A. Thomas, D. A. Taylor, D. L. Call, and D. I. Padberg, "Consumer Reaction to Nutrition Information of Food Product Labels," Search, Agriculture 2 (1972), 1-27.
- J. P. Liefeld, F.H.C. Edgecombe and Linda White, "Demographic Characteristics of Canadian Consumer Complainers" Journal of Consumer Affairs, 9 (Summer, 1975), 73-80.
- Fabian Linden, "The Consumer's View of Value Received--1974," The Conference Board Record, 11 (November, 1974), 48-53.
- Sidney Margolius, "The Consumer's Real Needs," Journal of Consumer Affairs, 9 (Winter, 1975), 129-138.
- Barbara D. Nichols and Rachel Dardis, "Consumer Satisfaction with Home Furnishings," Textile Chemist and Colorist, 5 (September, 1973), 23-27.
- Anita B. Pfaff, "The Quality of Consumption," in B. Strumpel, ed., Economic Means for Human Needs (Ann Arbor: Survey Research Center, Institute for Social Research, University of Michigan, 1976).
- Joseph T. Plummer, "Life Style, Social, and Economic Trends Influencing Consumer Satisfaction," in H. Keith Hunt, ed., Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction (Cambridge, Massachusetts: Marketing Science Institute, 1977).
- Daniel Powderly, "Care Labeling: Success or Failure?" Textile Chemist and Colorist, 8 (September, 1976), 41-43.
- Mary Shaw Ryan, Clothing: A Study in Human Behavior (New York: Holt, Rinehart and Winston, Inc., 1966).
- Sentry Insurance, Consumerism at the Crossroads (Stevens Point, WI: Sentry Insurance Co., 1977).
- George B. Sproles, "Clothing Orientations of Adult Women in Indiana," Research Bulletin #944, Agricultural Experiment Station, Purdue University, 1977 (in press).
- George B. Sproles, Loren V. Geistfeld and Suzanne B. Badenhop, "An Experimental Simulation on the Efficiency of Consumer Performance," Proceedings, American Council on Consumer Interests, 1977 (in press).
- Lynn B. Steiniger and Rachel Dardis, "Consumers' Textile Complaints," Textile Chemist and Colorist, 3 (July, 1971) 33-37.
- Mary Sue Stover, "Durable Press Shirts: Consumer Satisfaction and Effects of Wear and Laundering," Journal of Home Economics, 64 (March, 1972), 28-32.
- John E. Swan and Linda Jones Combs, "Product Performance and Consumer Satisfaction: A New Concept," Journal of Marketing, 40 (April, 1976), 25-33.
- Rex H. Warland, Robert O. Herrmann and Jane Willits, "Dissatisfied Consumers: Who Gets Upset and Who Takes Action," Journal of Consumer Affairs, 9 (Winter, 1975), 148-163.

THE PROCESS OF SETTING STANDARDS AND SOME
CURRENT ISSUES IN STANDARDS

L. J. Sharman
Consumer Product Safety Commission

Abstract

Each year an estimated 21 million Americans are injured in accidents in and around the home. The Consumer Product Safety Commission was formed by Public Law 92-573 with the objective of providing safer consumer products. Using the Flammable Fabrics Act as a basis, textile flammability standards now in effect are outlined and factors relating to their development discussed. Current Commission activities are reviewed.

Each year an estimated 21 million Americans are injured in accidents, including flammability accidents, in and around the home; 100,000 are permanently disabled; and 26,000 are killed. Many of these are product related. The economic and social cost to the Nation has been projected at \$5.5 billion. The urgent need for Federal Regulations to insure safe Consumer products was clearly recognized by Congress when it passed the Consumer Product Safety Act. That act, public law 92-573 was signed into Law by the President, October 1972.

The purposes of the Consumer Product Safety Act stated in the law, are:

- To protect the public against unreasonable risks of injury associated with consumer products;
- To assist consumers to evaluate the comparative safety of consumer products;
- To develop uniform safety standards for consumer products and to minimize conflicting state and local regulations; and
- To promote research and investigation into the causes and prevention of product-related deaths, illnesses and injuries.

To fulfill these purposes, the Consumer Product Safety Act created a new independent federal regulatory agency, the Consumer Product Safety Commission. This Commission headed by five Commissioners came into being in May 1973.

The Consumer Product Safety Act is a new and broader approach to product safety. The act also provided that many consumer product safety regulatory functions originally assigned to other federal agencies were transferred to the Consumer Product Safety Commission. The acts transferred were the Flammable Fabrics Act, the Federal Hazardous Substance Act, the Poison Prevention Packaging Act and the Refrigerator Safety Act.

Using these Acts, the Commission's primary goal is to reduce or eliminate unreasonable risks of injuries associated with consumer products. The Commission was granted broad authority to issue and enforce safety standards for more than 10,000 consumer products. Firearms, Food, Drugs, Cosmetics, Medical Devices, Pesticides, motor vehicles, airplanes and boats are exempted from the Commission's authority under the Consumer Product Safety Act.

When the Commission decides that a safety action is necessary, this action can take a variety of forms from educational programs, or voluntary standards at one end to the setting of mandatory safety standards or banning the product on the other. Imported products are covered by the act. The Commission also has the authority to order public notice, repair, as well as refund and replacement of defective products. The choice of repair, refund, or replacement is left to the manufacturer, distributor or retailer.

Where a mandatory consumer product safety rule is required, it may be a standard covering the performance, composition, contents, design, construction, finish or packaging of a product; the labelling of the product to warn or instruct the consumer as to the products' use; or it may be a rule totally banning the sale of the product. Additionally, rules and regulations may be prescribed for record keeping and other matters related to compliance with a standard or other regulation, and for the general administration of the act.

Of the various acts administered by the Commission, the Consumer Product Safety Act, the Flammable Fabrics Act and the Federal Hazardous Substances Act recognize flammability explicitly as a possible hazard associated with the products covered by the Acts and therefore, as a basis for regulations under the acts. Today, I plan to deal only with the Act which most clearly affects textiles, the Flammable Fabrics Act.

Mandatory Activities on Flammable Fabrics started with the passage by Congress, in 1953, of the Flammable Fabrics Act. The act referenced two voluntary standards, CS 191-53 and CS 192-53 making them mandatory flammability standards. In the case of CS 192-53, only the flammability parts were mandated.

The primary purpose of CS 191-53 was to remove the notorious "torch sweaters" and the highly flammable "Cowboy" play suits from the market. In this regard the standard was successful. As years went by, however, it was clearly evident that this standard was not preventing a significant number of serious flammability accidents. In 1976 the act was amended. The amendments broadened the coverage of the Act to wearing apparel and interior furnishings and gave the Secretary of Commerce the right and responsibility to set new fabric flammability standards where needed. This responsibility was transferred to CPSC by the Consumer Product Safety Act.

Since 1967, five new flammability standards have been promulgated. Two of these cover the flammability of Carpets and Rugs, FF 1-70 (16 CFR 1630) Standard for the Surface Flammability of Carpets and Rugs and FF 2-70 (16 CFR 1631) Standard for the Surface Flammability of small carpets and rugs. Two standards cover Children's Sleepwear, FF 3-71 (16 CFR 1615) Standard for the Flammability of Children's Sleepwear; sizes 0 through 6X and FF 5-74 (16 CFR 1616) Standard for the Flammability of Children's Sleepwear; sizes 7 through 14. The fifth, FF 4-72 (16 CFR 1632) covers the flammability of mattresses.

For all of the above standards the formal development

procedure involved three steps:

(1) Simultaneous publication in the FEDERAL REGISTER of advance notice of finding that a flammability standard may be needed and a notice instituting proceedings for the development of an appropriate flammability standard.

(2) Publication in the FEDERAL REGISTER of the proposed flammability standard for public comment.

(3) Publication in the FEDERAL REGISTER of the adopted flammability standard or notice terminating or suspending the proceeding to establish the standard.

This procedure was simplified by the Commission in 1976 by combining steps (1) and (2) above.

While the formal procedure may involve only 3 or 2 stages, the technical development of a standard involves many stages. Any standard is the result of composite effort by interested and involved people from industry, government and consumer groups. Almost every stage of a standard is a best compromise between the ideal situation and what is most practical and reasonable.

Using the Flammable Fabrics Act as a basis I will attempt to illustrate some of these points.

A standard starts with the determination that some kind of rule or regulation is needed. In the case of flammable fabrics standards these may be based on accident data or laboratory research. The five flammable fabrics standards promulgated to date were based on a combination of injury data and research. Once the potential need for a standard on a product has been determined, then, in-depth accident reports are reviewed to determine the causative factors. In the case of children's sleepwear, for example, the primary ignition source was an open flame while with mattresses the primary ignition source was a cigarette accidentally dropped on the mattress and left to burn. Using the major causative factors a rough test method is set up. Then the detailed work begins of checking possible ramifications and refining the test method.

The problem in the case of carpets and rugs was that of a small ignition source, such as a burning ember from a fireplace, igniting the carpet. This small fire wends its way across the carpet until it encounters other combustible products and ignites the room. For the performance test in the standard, a timed burning tablet is burned on each of eight, oven-dried, horizontal carpet specimens, in a draft protected chamber. A carpet fails the standard if more than one specimen chars 3 inches from its center (specifically to within 1 inch of the edge of an 8 inch hole in the specimen flattening frame).

One of the problems in developing this test method was to find a suitable ignition source which would burn repeatably and reproducibly. An extensive search located a tablet used by doctors in carrying out urine sugar tests. This methenamine tablet burns consistently 110 to 120 seconds. This tablet commonly referred to as the "pill" was incorporated into the test method.

When the carpet and rug work was first undertaken a test method was developed applicable to all types of carpets. A review of accident data and information supplied by the trade indicated that small carpets and rugs, such as bathroom mats, were not as great a hazard as large carpets and rugs. Two standards were, therefore, promulgated. One covers large carpets and rugs (a large carpet or rug is defined as having one dimension greater than 6 feet and a surface area greater than 24 square feet); and the other, small carpets and rugs (a small carpet or rug is defined as having no dimension greater than 6 feet and a surface area not greater than 24 square feet). The test method and acceptance criteria are the

same for both standards. The difference is that small carpets and rugs which fail the standard may be sold provided they are suitably labelled.

The first of the children's sleepwear standards was developed to protect children up to six years of age. Since this is a group of people who, to a large extent, do not have the capability of helping themselves the guiding criterion in developing the standard was - the standard must provide maximum protection for the children involved.

In depth accident reports indicated the primary problem with children's sleepwear to be open flame ignition. A vertical open flame ignition test method was, therefore, selected. In the performance test, oven dried 3.5 X 10 inch specimens are suspended vertically in a holder in a prescribed cabinet and subjected to a small flame for 3 seconds, along the specimens bottom edge. Acceptance criteria are based on char length and residual flame time (burning time of material which falls away from the specimen and continues to burn on the floor of the test cabinet). Tests are required on fabric, seams and trim. The standard includes a sampling plan specifying the required frequency of testing.

During the development of this standard a wide range of factors related to the test method were investigated. One of the most interesting involved the ignition time of the fabric. Prior to the development of the first children's sleepwear standard, most vertical flammability tests involved subjecting the test specimen to a small flame for 12 seconds. During the development of the standard the effect of varying the time the specimen was subjected to the flame was investigated. It was found that with some fabrics the specimen did not ignite with 12 seconds exposure but did ignite with 2 or 3 seconds exposure. Further it was found that a burning specimen could be extinguished by again subjecting the specimen to the igniting flame. There are a number of theories as to why this phenomena occurs. One of the commonly held ones is that of oxygen depletion. Holding the flame under the specimen for a long period of time creates air currents which remove oxygen from the fabric and prevent its ignition or puts the fire out if the specimen is already burning. Based on this work, 3 seconds was selected as the time of flame exposure for the test specimens.

FF 3-71 was the first flammability standard to include sampling plans. Without a sampling plan, a standard does not state how frequently the product must be tested or if a failure occurs what quantity of the product is involved in the failure. Here, as in the test method, a sampling plan is a series of compromises between the ideal and what is practical. One compromise that still bothers some of the purists is that of fabric sample selection. Working with representatives of the textile industry, 5000 linear yards of treated fabric were selected as a fabric production unit for testing purposes. The next question was how to select a representative sample from this unit. The ideal, of course would be to randomly select the test specimens throughout the 5000 yards. This however is not very practical for many reasons. With most textile operations the worst part of the fabric occurs at the beginning or end of the fabric roll. On this basis it was decided that the specimens should be selected from the ends of the 5000 yard production unit.

The second children's sleepwear standard FF 5-74 covers children in the 6 to 12 age range. The test method in this standard is the same as that of FF 3-71. The primary difference is in acceptance criteria. FF 3-71 acceptance criteria involve char length and residual flame time. FF 5-74 involves only char length. Residual flame time was not included in this standard on the basis that the standard covered sleepwear used by older child-

ren who would be capable of helping themselves in the case of a fabric fire.

The last flammable fabrics standard in force today is that for mattresses. Here the primary problem is ignition of the mattress by a lit cigarette accidentally dropped on the mattress. This results in a smoldering fire and the rapid generation of a lethal atmosphere.

In the Standard's performance test, lighted cigarettes are placed on temperature and humidity conditioned mattresses, both on the bare mattress surface and between two layers of sheeting. The cigarettes are placed on smooth surfaces, along the taped edges and in depressions formed by quilting or tufting. A mattress must not ignite from any of 18 cigarettes.

One of the questions raised in developing this test method was what cigarette to use in the test. A wide range of commercial cigarettes was evaluated. For the most part, the burning temperatures of the cigarettes were similar. It was found, however, that unfiltered cigarettes burned much hotter at the butt end than did filter cigarettes. For this reason unfiltered cigarettes were selected for the test method.

Another question area investigated was that of the permanency of any treatment used to make the mattresses resistant to ignition by cigarettes. To check this, the initial test method included a procedure for wetting down the test mattress, drying it and then testing. This procedure was very cumbersome, time consuming and costly. Working with a consultant, who had been employed by industry to develop mattress constructions resistant to cigarette ignition, it was established that there was very little likelihood of a permanency problem. The wetting down requirement was, therefore, dropped from the final test method.

The Commission is also actively working on a number of other areas related to fabric flammability. One of these is the flammability of upholstered furniture. The problem here is similar to that with mattresses - a lit cigarette accidentally dropped on a piece of upholstered furniture, ignition of the furniture item and the relatively rapid generation of a lethal atmosphere. The National Bureau of Standards under contract to the Consumer Product Safety Commission have developed a test method for upholstered furniture. This test is similar to that for mattresses in that it involves determining the resistance of the furniture construction to ignition by cigarettes. The NBS test procedure involves testing of and classifying fabrics for use on upholstered furniture and testing of mock-ups constructed in the same way and of the same materials as would be used in the actual furniture.

The fabric test involves determining the char length of the fabric when exposed to cigarettes over standard substrates. Depending on the results obtained, fabrics are classified as A, B, C or D with A being the most ignition resistant and D the least. None of the fabric classifications represent "failure" under the proposed test method. However, to make upholstered furniture items which will meet the acceptance criteria, fabrics in one classification will in general, have to be used over different upholstery materials and/or methods of construction than will fabrics in other classifications.

To determine the resistance to cigarette ignition of the furniture construction on which the fabric may be used, a mock-up of specific surface locations utilizing the same filling materials and method of construction as would be used for the actual furniture is prepared. Fabrics specified in the test method are used to cover the mock-up. After conditioning, the mock-ups are tested by burning cigarettes in horizontal crevices where seat cushions and vertical panels meet; on seat cushion surface locations including smooth surface, quilt,

tuft, and welt edge; on top surfaces of armrests, on top surfaces of backs and on top surfaces of loose support systems. The number and types of locations to be tested will depend upon the construction and design of the production furniture to be qualified. Three cigarettes are burned at each of the specified locations. If two or more ignitions occur the construction is not acceptable. If one ignition occurs then three additional cigarettes are burned on replicate mock-ups at the ignited location. If none of these results in an ignition, the construction is accepted. This repeat test allows for the detection of fluke ignitions without significantly affecting the safety being provided the consumer. When all mock-up surface locations have been shown to meet the acceptance criteria for a specific fabric class, then furniture may be manufactured using the construction, and any fabric in the specified class or any higher class. This does not apply to class D fabrics where each fabric must be tested with its proposed construction.

In June, the CPSC Commissioners formally considered the upholstered furniture flammability problem. The Commissioners agreed in principal that there is a flammability problem with upholstered furniture and that the Commission's objective should be the production by industry of furniture that would not ignite when a cigarette is allowed to burn on it. Further, they agreed in principal that the NBS test method for cigarette ignition resistance of upholstered furniture was reasonable and repeatable. They expressed concern, however, that the proposed application of the test method (such as frequency of testing) may be too burdensome and restrictive to the industry. CPSC staff were directed to propose alternative test applications for Commission consideration as well as alternative methods of implementing the standard. This work is in progress.

A second area under investigation is that of wearing apparel. Standards have been promulgated for children's sleepwear. Are other standards for specific apparel items needed? CS 191-53 has not prevented many apparel flammability accidents. Is a new general wearing apparel standard needed? Towards this latter end, NBS at the Commission's request has developed a potential test method. This test involves measuring the ignition time and the heat transfer rate of a fabric. Depending on the results obtained, the fabric is classified 1 through 4. The wearing apparel items in which fabric may be used depends on its fabric classification. This test method has become known as the "Mushroom Apparel Flammability Test (MAFT)".

An alternative approach would be to concentrate on only those areas of wearing apparel where injury data indicates the hazard to be high and to develop standards for these. Available data indicate the primary hazard areas to be nightgowns, (those not now covered), Robes and Housecoats (Females only), Pajamas (Male and Female), Dresses, Shirts (Male) and Pants (Male). One potential approach for specified wearing apparel would be a modified form of the children's sleepwear standard FF 5-74. At this time, the Commission has not made a decision on what further action to take. A concept briefing package covering the overall apparel flammability question is being prepared for the Commissioners with a request for guidance to the staff on the approach to be taken.

Lastly, I am sure you are all aware of the tris (2,3 dibromopropyl) phosphate (TBPP) controversy earlier this year. TBPP was used as a flame retardant on a significant portion of the children's sleepwear produced. Industry have requested that the sleepwear standard be amended to permit the use of more untreated fabrics. A discussion paper has been prepared and is now under consideration by the Commissioners covering the aspects of the standard which industry has requested consideration for amendment. These include the need for the residual flame time criteria, the need to test seams and trim, procedure for

conditioning specimens, exclusion of the 0-1 age group, exclusion of tight fitting garments and modification of the sampling plan.

Early in this talk I stated that any standard is the composite effort of all interested and involved people. Through the rest of the talk, I have tried to exemplify these statements. A further exemplification is that such people are on today's session. Rachel Dardis who spoke to you earlier is a member of the National Advisory Committee for the Flammable Fabrics Act. The American Textile Manufacturers Institute along with other textile organizations such as the American Apparel Manufacturers Association and the Man-Made Fiber Producers Association have been a great deal of help in the past and will be in the future. We need such help from all of you.

HUMAN VALUES: A HISTORICAL AND INTERDISCIPLINARY ANALYSIS

C. Joseph Clawson, University of Southern California
Donald E. Vinson, University of Southern California

Abstract

... Values are so inextricably woven into our language, thought and behavior patterns that they have fascinated philosophers for millennia. Yet they have proved so "quick-silvery" and complex that, despite their decisive role in human motivation, we remain desperately ignorant of the laws that govern them. (Toffler, 1969)

One of the most curious anomalies in the field of consumer research has been the lack of attention devoted to human values. This is particularly paradoxical considering that most serious students of human behavior have argued for centuries that values play an important role in personal, social and cultural activity. Where the term "values" has been employed, it has typically been confused with the concept of product attributes by attitude researchers. While a few contemporary studies have emerged suggesting significant correlation between values and consumer behavior (Scott and Lamont, 1972; Vinson and Munson, 1976; Henry, 1976; Vinson, Scott and Lamont, 1977), consumer researchers have, for the most part, adopted a hands-off attitude toward the value construct. In dramatic contrast, marketing practitioners have become very sophisticated in developing advertising and mass communication appeals involving allusions to important human values. In fact, we would argue that marketing appeals have traditionally been highly value oriented. Apparently, marketing practitioners understand what most marketing academics fail to recognize -- human values play an important role in consumer behavior.

What could explain this disparity between what marketers do and what we in academia study? One plausible explanation may be that academics have concentrated their attention and technical skills almost entirely on consumers' more rational beliefs with respect to tangible product attributes and marketing characteristics -- such realities as color, price, taste and availability. A contrasting but much less popularly held view is that practitioners may have forged years ahead of our scientific decision theories. While academic researchers argue about A₁'s and B₁'s and report low R²'s from attitude studies, practitioners are appealing to deep, emotion-laden values and providing powerful stimulation to sales.

In order to emphasize that the analysis of human values is not just another new fad in consumer research, or a minor variation on product attribute studies, we shall present a historical review of the value concept from antiquity to the present day. However, the subject of values has been of intense interest through much of recorded history and has generated a vast literature. Consequently, our discussion must be narrowed by two rules. First, we shall be concerned only with analyzing subjective values that consumers actually hold. These are distinct from objective or true values, as might be determined by scientific facts or religious insights. Second, we shall focus on the values that objects, such as goods and services, have to consumers, but only insofar as market value can be related to specific human values that people themselves hold. This seemingly innocent restriction eliminates almost, but not quite, all of economic theory of consumer demand

for the past two hundred years. What is often called value theory in economics is primarily concerned with market value, not human values.

Historical Perspectives

Initially, most of the thinking on market value and consumer values was contributed by the philosophers and religious leaders of Europe. During the late 1700's, there began to emerge various branches of philosophy which we now call the "social sciences."

Philosophy

Since the days of the ancient Greeks, philosophers have been concerned with values on a rather tangential basis. References to values were unavoidable as Aristotle, Kant, and others discussed aesthetics, or as Plato, Hobbes, and Rousseau deliberated over the problems of government and citizen responsibility. But, as Wekmeister points out (1967), no general theory of values was developed or enunciated by any of these thinkers.

Some ancient insights into the nature and importance of wants, needs, or values as the motivators of consumer demand are worthy of note. Aristotle argued that man obtains greater pleasure from an object when it is his own, for the love of self is a feeling implanted by nature and not given in vain. There is also great pleasure in doing a kindness or service to friends or guests or companions, which is facilitated by possessions (Haney, 1936). Thus pride of possessions and hospitality were identified as what we would now call consumer values.

Buridan (1300-1358) followed Aristotle in stating that the measure of value in goods is to be found in the satisfaction of wants, greater need resulting in higher value. Samuel Pufendorf (1632-1694) asserted that the price or value of any action or thing depends on its fitness to obtain, directly or indirectly, the necessities, conveniences, or pleasures of human life. Nicholas Barbon in 1680 refuted the notion that economic ends are necessarily tangible, or that pleasures and pains have to be physical (*Ibid.*). In parallel thought, Galiani in 1750 stressed the desire for social distinction, in the form of rank, titles, honors, nobility, and authority, which he held to be stronger than the desire for luxuries, which in turn was stronger than the desire of the hungry for food (Georgescu-Roegen, 1968).

Jeremy Bentham (1748-1832), a non-practising lawyer, was a major contributor to the hedonistic and instrumentalistic interpretations of behavior. His Introduction to the Principles of Morals and Legislation (Patterson, 1932) revealed his conviction that cultural values as well as all other instrumental acts are ultimately learned through reinforcement or inhibition, which he called the principle of utility. According to him, utility is that property in any object which produces benefit, advantage, pleasure, good, or happiness, or prevents mischief, pain, evil, or unhappiness to the party making the decision. One of Bentham's less well known works, A Table of the Springs of Action (1815), foreshadowed Kurt Lewin's psychological field theory as well as later compensatory multiattribute models. He

observed that multiple, conflicting motives are often simultaneously at work, with the dominant motive or group of motives determining the final decision (McReynolds, 1969). Bentham's fourteen springs of action foreshadowed by more than a century some of the great value lists to be discussed later. He included taste, sexual appetite, wealth, power, curiosity, amity, reputation, religion, sympathy, antipathy, self-interest, and the pains of labor and physical suffering.

The question of values first began to emerge as a separate and unique issue in Germany during the early 1800's. Indeed, intense scholarly debate as to the source, nature, and even existence of things called values signaled the acknowledgment of axiology (the study of value) as a legitimate and fruitful area of academic inquiry. The basic problem confronting the study of values from a philosophical perspective relates to the axiological question of whether or not values are subjective or objective. Those proposing the "objectivity" of values are in essence suggesting that values are independent of a subject or a valuing consciousness. Conversely, the "subjective" position holds that values owe their existence, their validity, to the subject who does the valuing (Fronzizi, 1963). As a case in point, one could evaluate an individual's appreciation of a beautiful painting as either the recognition of certain values existing and associated with the particular work of art, or as the person's personal preference or liking of the objet d'art which gives it its value.

Meinong was the first subjectivist to address himself to the problem of the nature of values. Something has value, he said, "when it pleases us, and to the extent that it pleases us" (*Ibid.*, p. 34). In a similar vein, Ehrenfels believed, "Things which we desire or covet are valuable, and they are so because we desire and covet them" (*Ibid.*, p. 37). Supporting the subjective position, Bertrand Russell said, "The chief ground for adopting the subjective view of values is the complete impossibility of finding any arguments to prove that this or that has intrinsic value." It is "our desires which confer value" (Russell, 1925).

The objective approach to values was championed by Hartmann and Scheler. These two scholars believed that the study of values represented the investigation of a legitimate scientific field, unlike Russell who held that questions which refer to values were outside the realm of science. Hartmann approached the concept of value in strictly logical terms; to him, the value predicate is a property of concepts rather than objects (Hartmann, 1967). Scheler argued that, if all values were relative to life, as the subjectivists suggest, then life itself could not possibly have any value. Value independence implies their immutability; they are not dependent upon any act, regardless of its nature, be it historical, social, biological or purely individual. It is only knowledge of values which is relative, not values themselves (Fronzizi, p. 82). In addition to values being independent of their corresponding carriers, Scheler held that they exist in a hierarchical order. This order is expressed by "preferring" or "deferring"; we prefer one course of action, one good, one mode of behavior to another as we prefer "a given value to another, regardless of the carrier" (*Ibid.*, p. 84).

Those value philosophers of the objective persuasion acknowledge that valuation is indeed subjective; however, one must distinguish between valuation and value. On this point, Fronzizi states that value is prior to valuation -- if there were no values, what would people evaluate (*Ibid.*, p. 15)? Values then are those things which are evaluated in the valuation process.

Modern social thinkers have generally embraced the objectivist position and regard the emergence of values as a social-cultural process. Nietzsche (1844-1900) inter-

preted the dynamic character of history as the continuous creation and annihilation of values. Men create values which stabilize cultures, albeit temporarily, until these values are subsequently replaced by another set which allow the upsurge of a new human culture (*Ibid.*, p. 33).

Behavioral Sciences

Anthropology. Anthropologists have contributed enormously to the study of human values through their examination of cultural patterns and life styles. Clyde Kluckhohn shifted anthropological thinking away from the view that all cultures are relative. Despite wide differences in customs, there are certain fundamental human values common to all of the diverse cultures of the world (Kluckhohn, 1951). Further, Kluckhohn devoted considerable effort to the classification of values as well as to the distinguishing dimensions of values. Expanding his work, Florence Kluckhohn and Strodtbeck developed a five-part typology of value orientations and employed it to build a theory of cultural variation. They concluded that society requires conformity if man is to be a social animal and that societies create the value system to which men must conform (Kluckhohn and Strodtbeck, 1961).

Another anthropologist interested in values was Dorothy Lee, who studied the Arapesh of New Guinea in their food gathering activities. She discovered that the structuring of their food-getting procedures was appallingly inefficient from the point of view of food production, but it was eminently efficient from the point of view of infusing [values] into their lives (Lee, 1959). As Ayres so appropriately states, ". . . it is perfectly true that every culture includes a considerable admixture of irrationality -- that values exist and are cherished with the highest degree of intensity which owe virtually nothing to science but are virtually pure expressions of community sentiments and attitudes fortified by tribal legends and ceremonials of every kind" (Ayers, 1950).

Sociology. Ferdinand Tonnies regarded values as a critical element in his classic distinction between community and society -- between those systems which emphasized Gemeinschaft (primary, small, traditional, integrated values) and those which stressed Gesellschaft (impersonal, secondary, large, socially differentiated values) (Tonnies, 1957). In a similar vein, Durkheim described his "collective consciousness" as a system of values and beliefs held in common by the members of a society that define what their mutual relations ought to be (Durkheim, 1960). In the case of mechanical solidarity, the group's values are not clearly differentiated from the norms through which they are implemented; in organic solidarity, the norms have independent salience.

The study of values was considered the core of sociological inquiry for Becker (1941). He considered values as indispensable in analyzing judgments of right and wrong, good and bad, superiority and inferiority, usefulness and uselessness, and in determining the means and ends of human action. In agreement, Blau (1964) suggests that a prime determinant of social conduct is the institutionalized system of values in a society. Values define group identity, common standards of morality and achievement and legitimize governing authority. While the sociology literature contains voluminous references to values, some of the more significant contributions involve the following areas of sociological inquiry: measurement and quantitative techniques (Lundberg, 1941); "pattern variable" analysis (Parsons, 1969); deviant behavior (Parsons, 1951); content mobility in education (Turner, 1961); social differentiation (Sorokin, 1947); social exchange (Blau, *op. cit.*); socialization (Brim, 1966); and the theory of values, value system organization and the role of values in social behavior (Rokeach, 1968a, 1968b, 1968-69).

Psychology. Psychologists have generally focused on more narrowly circumscribed constructs such as attitudes, motives, valences, and cathexes. However, in the 1920's, Thomas and Znaniecki recognized the importance of values and defined them as ". . . objective, social elements which impose themselves upon the individual as a given and provoke his reaction" (Thomas and Znaniecki, 1927).

Another early landmark in psychological analysis of values was the publication of the English translation of Spranger's *Types of Men* (Spranger, 1928). Believing that values are reflected in all behavior, Spranger constructed a value system typology which purported to classify men according to six basic personality characteristics. Values could be clustered into systems representing types of men whose personality manifestations tended to be theoretical, economic, aesthetic, social, political, or religious. He did, however, admit that personality did not depend exclusively upon any one of these value orientations in total; he conceded that all six orientations probably exerted influence to varying degrees. Nevertheless, it was his contention that the individual's predominant value orientation structured his personality.

A renowned empirical investigation emerging from the Spranger tradition was undertaken by Allport and Vernon during the early 1930's (Allport and Vernon, 1931). Their study employed an instrument designed to test the relative strength of the values associated with Spranger's six personality types. In replications over a large number of heterogeneous subjects, they reported reliability measures of over .72. Corroborating these findings, Pinter (1933) and Evans (1952) indicate the validity of the "Types of Man" approach to the study of values and support the notion that values are socially learned.

A contemporary of Allport's, also teaching at Harvard in the late 1930's, was Henry A. Murray. As part of a team of clinical investigators, he developed what is one of the most widely used list of needs today (Murray, 1938). The list contains 28 secondary or psychogenic needs and 12 viscerogenic needs. In an easily overlooked passage, Murray clearly pointed out that needs are not identical with values: ". . . From one point of view, the important thing is not whether a subject has a need for Achievement or for Affiliation or for Rejection, but rather what it is he wishes to achieve, affiliate himself with, or reject."

By 1942, Kurt Lewin was conducting research using the equivalent of a simple linear compensatory version of the expectancy-value formula and applied it to a mixture of four product attributes and personal values. These "frames of reference" consisted of expense, health, taste, and status. He asked three questions: (1) What are the values for this group? (2) What is the relative weight of each value? (3) How are specific foods linked with certain values? (Lewin, 1943).

In the mid 1950's Maslow advanced his hierarchical groupings of needs into those that were concerned with safety, security, love, self-esteem, and self-actualization (Maslow, 1954). Morris (Handy, 1970) discussed value theory in considerable detail and developed a list of thirteen "ways to live" which were conceptions of the good life that respondents were asked to rate as to degree of liking. Rosenberg (1956) suggested that values and needs are pervasive in the individual's cognitive structure and attempted to measure the importance of a given value or set of values relative to an attitude object. Thurstone then proposed that values could be measured and, with the aid of the comparative judgment technique, developed a scale of "moral" values (Thurstone, 1954). W. A. Scott described a method for measuring the values and ideologies of a culture (Scott,

1959). Woodruff and Divesta measured the values of college students and showed their relationship to attitudes (Woodruff and Divesta, 1948).

One of the most influential of modern contributors to the psychological theory of personal values is Milton Rokeach. According to Rokeach, adults possess many thousands of attitudes toward specific objects and situations but relatively few terminal and instrumental values -- perhaps only several dozen instrumental values and a few "handfuls" of terminal values (*Ibid.*, 162). This difference in number suggests to him that values and attitudes are connected in a hierarchical system. This cognitively connected, hierarchical structure is arranged along a central-peripheral dimension and in a psychological sense is internally consistent. Any change in one part of the value-attitude system will affect the other parts. However, the most central or salient elements are more resistant to change and, when they do undergo a restructuring, they exert greater influence upon the other components of the system.

The reliability of the Rokeach paradigm has been verified under many replications (Robinson and Shaver, 1971) and Rokeach has found that various combinations of terminal and instrumental values significantly differentiate among people with extremely diverse social and economic status characteristics (Rokeach, 1968-69). Unlike most cognitive approaches to social psychology, his paradigm focuses upon values rather than attitudes. According to Rokeach, values are culturally derived, and define, maintain and regulate the visible social structure; they give it meaning, stability and cohesion. "To say that a person 'has a value' is to say that he has an enduring belief that a specific mode of conduct or end-state of existence is personally and socially preferable to alternative modes of conduct or end-states of existence (Rokeach, 1968, p. 167). Instrumental values relate to modes of conduct and terminal values have to do with preferred end-states of existence. The eighteen terminal and eighteen instrumental values developed by Rokeach had test-retest reliabilities in the .70's after seven weeks (*Ibid.*, p. 167).

Like values, attitudes are also based upon beliefs. But rather than a single belief, they represent the focusing of several beliefs upon an object or situation and predispose the individual to respond in some preferential manner.

For purposes of research in consumer behavior, Rokeach's excellent paradigm contains a few limitations. Three should be noted. One flaw is the use of rank orderings, which, for all its speed and simplicity, is less informative than interval or ratio scaling would be. Equally attractive values are forced into separate rankings. Wide gaps in preference are treated as no different from minuscule gaps.

A second problem is with the instructions, which bias the rankings in favor of deprived values and against satiated values. Saying to a respondent that his task is to arrange the value statements in order of their importance to him, as guiding principles in his life (Rokeach, 1973, p. 358), does not allow him to distinguish between permanent esteem of the value and situational need for the value. For instance, the lowest income respondents rate "a comfortable life" and "clean" relatively high, while wealthy respondents rate them quite low. The low income respondents presumably need, or lack, these conditions, the high income respondents already possess them, but both groups could esteem them equally. Similarly, black Americans need and rate equality very high, while white Americans need it less and rate it as only moderately important (*Ibid.*, p. 393). It is still not known whether there is a real difference between whites and blacks in permanent esteem of equality.

The third drawback is that Rokeach's lists of 18 instrumental values and 18 terminal values appear to omit a substantial number of other values that are held by substantial portions of the populace. They include, for example, Life, Physical Energy, Physical Strength, Physical Attractiveness, Youthfulness, Family, Enjoyment, Power, Leadership, Courage, Individualism, Conformity, Mental Energy, Practicality, and Efficiency. Such values are often appealed to in advertising.

The final psychologist making a distinguished contribution to the theory of human values, applicable to consumer behavior, is McGuire. His most complete analysis is published as a chapter in a book on gratifications obtained through mass communications (McGuire, 1974). His classifications of motives, he believes, "must sensibly be expanded to include lists of instincts, . . . emotions, . . . values, . . . temperaments, . . . and motivation-connoting personality dimensions" (*Ibid.*, p. 170; italics added). The impressive and useful result is a table consisting of 16 classifications of motives.

Value Measurements in Behavioral Sciences. The fertile generation of value-related concepts by anthropologists, sociologists, and psychologists has been paralleled by a stream of value-measurement scales. Twelve scales which may be of use to consumer researchers are reproduced in a single document (Robinson and Shaver, *op. cit.*), with references given after each scale.

Marketing. The first academic contributions to formal marketing thought on consumer values were made by Harvard's Melvin Copeland in 1924. After analyzing the appeals used in 936 advertisements in American magazines published during 1923, he concluded that all the buying motives to which they appealed could be classified on two dimensions. First, when sorted according to the type of decision that was to be made, motives could be sorted into "primary" motives affecting choice of product class, "selective" motives influencing brand choice, and "patronage" motives determining store choice (Copeland, 1924).

The second dimension was considered by Copeland as more basic. He said that motives could also be sorted according to elements entering into the decisions; this resulted in "emotional" motives being distinguished from "rational" motives (*Ibid.*, p. 162).

According to Copeland, the 23 emotional motives had their origin in human instincts and emotions. They represented impulsive or unreasoning promptings to action, and could not be aroused by appeals to reason. Examples from his list are pride of personal appearance, social achievement, expression of artistic taste, romantic instinct, cleanliness, entertainment (*Ibid.*, p. 163-178). These motives closely correspond to the relatively central or global personal values which Rokeach has called instrumental and terminal values. So far as we have been able to ascertain, Copeland was the first marketing scholar to distinguish formally between what we consider as human values, as one category, and what are commonly referred to as product or marketing attributes, as a separate category.

A lengthier classified list of 128 values was developed by Clawson, a student of Allport's (Clawson, 1946, pp. 26-27, 271-274), and was based on content analysis of several hundred best-read advertisements and numerous propaganda materials. Six of the category titles corresponded to Spranger's Types of Men, but two additional categories were included. They consisted of 29 "physical" or biogenic values and seven "general" values.

Clawson also brought to the attention of marketers (1950) several of Lewin's concepts which would be helpful in linking the broad construct of human values with

the more specific marketing construct of buying motives. These concepts included the "achievement level," the "ideal level," and the expectation level, regarded as common to all values.

Even with the tremendous and growing impact of the behavioral sciences upon marketing in recent years, consumer values, as such, have been largely ignored by the discipline. Where they have received attention, they have been employed with almost poetic license and defined with common vernacular. As an example, in 1967, one entire section of the Proceedings of the American Marketing Association was devoted to "Value Systems in a High Level Economy" (Mayer, 1967). However, not one of the papers in this section presented a conceptual or operational definition of values, or related values to marketing theory or practice.

In the area of consumer behavior, values have usually tended to be equated with life style (Bem and Bem, 1973; Ireland and Besner, 1971; Mitchell, 1971), group norms (Kassarjian and Robertson, 1973), or social class differences (Levy, 1973). The life style association appears to be the most popular. In fact, Kelley and Lazer devote one chapter of their book, Managerial Marketing, to "Life Styles and Values" (1973). This title turns out to be somewhat of a misnomer in that little, if any, discussion pertains to values *per se*. Furthermore, nowhere in the chapter are the two constructs conceptually differentiated.

The recent infatuation with attitude models has also led to confusion concerning the role of values in consumer research. Borrowing freely from Rosenberg's value-expectancy conceptualization, marketing scholars have concentrated on the prediction of brand preference using product attributes instead of Rosenberg-type generalized values in their structural equations. Today, many marketing and consumer researchers mistakenly equate "values" with evaluation of product attributes. Although a complete analysis of expectancy-value models is beyond the intended scope of this paper, excellent reviews are available (see Wilkie and Pessemier, 1973; Jacoby, 1976; Hansen, 1977).

There are numerous indications of growing interest in personal values, as distinct from product and marketing attributes, in the marketing literature. One is the "Second Washington Social Indicators Conference to Explore the Response of Business to New Values" (Marketing News, 1973). In addition, there appears to be an emerging body of research devoted to the topic. Scott and Lamont (1972) have presented evidence that global values, product evaluative and descriptive beliefs, and domain-specific values are three separate yet cognitively related variables. Henry (1976) shows significant associations between automobile ownership and personal value-orientations along the Kluckhohn-Strodtbeck lines. Extending Rokeach's paradigm, Vinson, Scott and Lamont (1977) view values as existing at two distinct levels. The first level, referred to as "global" values, contains both instrumental and terminal values. The second level deals with values which refer to desired product attributes and market place transactions and behaviors. Termed "market-specific" or "domain-specific" values, they are cognitively separate but related, and are connected to global values as well as to descriptive and evaluative beliefs.

Vinson and Munson (1976) demonstrate that values could be productively employed as segmentation variables. Vinson, Munson and Nakanishi (1977) demonstrate the validity of the Value Survey for consumer research applications. Munson uses Rokeach's terminal and instrumental values to successfully differentiate consumers of culturally diverse nationalities (1977). And Vinson suggests that differential value orientations represent an important underlying dimension of consumer discontent

(1977).

Howard's new book devotes extensive space to consumer values (1977). He stresses instrumentalism (reflected in his frequent references and examples of lengthy means-end chains), and suggests that terminal values influence a person's choice criteria for a product class, his beliefs about product classes, his attitudes toward product classes, and finally his purchase of a particular product class (*Ibid.*, p. 98). A verbal example suggests that a consumer may perceive a means-end chain as consisting of the following links: meat-protein-energy-work-effectiveness-ambition (*Ibid.*, p. 116). Howard also cites numerous potential applications of consumer values in developing government and marketing policy decisions. However, his own stimulating analysis does not include any rigorous model of the structural elements, causes, or consequences of consumer values.

Conclusion

The study of consumer values shows many signs of becoming a challenging area for research in the years ahead. The most encouraging single indication is that consumer values have already been shown to have significant correlations with consumer behavior and attitudes. Given a solution to some of the present difficulties with the concept, values may prove to be one of the more powerful explanations of, and influences on, consumer behavior. They can perhaps equal or surpass the contributions of other major constructs including attitudes, product attributes, degree of deliberation, product classifications, and life styles. However, the untapped potential of consumer value analysis can be properly harnessed only if certain obstacles and problems can be removed. Our review reveals five major problems:

First of all, the concept of consumer values has not yet been clearly defined. Second, no comprehensive, non-duplicative list of values has been agreed upon by philosophers, behavioral scientists, economists, or marketers. Third, no generally accepted model is as yet available specifying the structural variables of an individual value, its latent or active interactions with other values, or its relations to other predispositions, environment, and behavior. Fourth, no standard method of measuring values has been adopted. Finally, the practical value of values has yet to be demonstrated or rejected conclusively, either in real-life settings or in laboratory settings. Insights into these applications would be needed by policy-makers in such fields as marketing, government, and other areas of planning and leadership.

Need for a Better Definition. Space permits us only to deal with the first and, in our opinion, most pressing of the five problems in value analysis. This is the problem of definition of a value. We may at least define what consumer values are, and what they are not, by indicating the major elements that such a definition should contain. These elements have been distilled from the work of many scholars and has been heavily influenced by our historical review.

(1) A consumer value is a belief. The first element of a proper definition recognizes that a value is a belief held by the consumer. It is not some objective "truth" that might have been tested and accepted by scientists, philosophers, religious leaders, economists, or other observers.

(2) A consumer value is person-centered. The requirement that human values be person-centered immediately reduces the number of eligible values. Person-centered values are those values explicitly connected in the consumer's mind with his own personal well-being or social acceptance, or with the well-being and social acceptance of other persons who are important to him.

This excludes object-centered values, which are those objects, events, activities, or even people in the consumer's environment that have no explicit effect, per se, on him or on significant others. The latter include (a) specific actions or action-sequences that vary from one situation to another, such as searching, buying, storing, processing, using, or repairing; (b) specific products, services, brands, stores, advertisements; (c) a product's or service's physical and marketing attributes; (d) its technical performance characteristics, and (e) its uses. For example, a breakfast drink's calorie content is a nonpersonal attribute. It is an object-value but not, per se, a human value. However, the nutritional effect on people of these calories is a human value, since it is explicitly person-centered. We believe that if we wish to understand the entire range of person-centered values that motivate consumer behavior, we should start, of course, with (1) Rokeach's modes of behavior (instrumental values) and end-states of existence (terminal values), but we would need to add (2) the wide range of biological and social states omitted by Rokeach, such as life, energy, health, sex, physical senses, power, fame, property; (3) relatively complex, long-term, recurrent patterns of personal skills, ideas, objects, and activities, such as good manners, professional competence, a particular religion, a particular political party, patriotism, individualism, physical attractiveness, life style, and personal hobbies, to name a few; and finally (4) specific, short-term personal benefits that are relevant in a wide variety of situations, including time saving, economy, ease, convenience of location, and the like.

The three extra categories we have suggested may of course be intertwined with one another and with the Rokeach-type values. Indeed, they may even be derived from his values. Nevertheless, they have a definite functional autonomy of their own and are often mentioned as goals by consumers and are appealed to by many advertisers, salesmen, political leaders, and other communicators.

(3) A consumer value is perceived as having an underlying positive worth, as an end in itself, for the consumer himself or for significant others, and is organized around a subjective ideal. The subjective worth may be called utility, importance, desirability, liking, approval, loyalty, or other terms that reflect the affective, or evaluative, aspect of every consumer value. The fact that a particular value has subjective worth to a consumer may be revealed in a variety of ways. It is revealed by overt behavior, emotion felt or revealed, and verbal expressions; by vicarious participation in the experiences and emotions of others; and by subconscious or concealed responses. Each value also has a subjectively ideal state, which is preferable to any higher or lower state. For instance, the right amount of generosity is regarded as preferable to the extreme of extravagance as well as to the opposite extreme of miserliness.

(4) A value is perceived as having present or eventual usefulness. Although a value is esteemed in part for its own sake, it is also regarded by the consumer as being an instrument for serving still deeper values. Even the terminal values, says Rokeach, are the conceptual tools and weapons that people use to maintain and enhance their self-esteem (Rokeach, 1973, p. 14), which McDougall called the master sentiment of self-regard (*Ibid.*, p. 14) and Maslow called self-actualization (*Ibid.*, p. 16, Maslow; 1959, p. 123). Even the self-concept would be useful as a criterion for evaluating other values, attitudes, and activities.

(5) A value is enduring. A value is neither completely stable nor completely unstable. If it were the former, it would be impossible for individual and social changes to take place; if the latter, neither personality nor

society would have any continuity (Rokeach, 1973, pp. 5-6). While values remain relatively stable, the needs and behavior derived from them can fluctuate widely. To illustrate, a high subjective worth placed on physical comfort is a relatively enduring value, and is not inconsistent with varying intensity of headaches, nor with purchasing aspirins on some days and not others. Researchers will need to take precautions to measure the relatively permanent, underlying esteem of values, undistorted by transitory needs, surpluses, threats, or inducements.

(6) A consumer value is widely held by many but not all members of a sub-culture, society, nation, or civilization. This element tends to eliminate rare hobbies and personal attachments of isolated individuals from possible classification as consumer values. However, unanimity is not required. An opposite or compromise position on any value may be adopted by a minority within any group being studied, sometimes with, sometimes without the tolerance of the majority. This recognizes the coexistence of rival ideologies, opposing political parties, criminal elements, moral deviates, terrorists, and others with nonconforming life styles within any social grouping. It does not preclude warfare among factions and nations, based partly on differential value systems and partly on differential degrees of deprivation.

(7) A major role of a consumer's value is that of standard or "criterion" he can use in the formulation of attitudes and guidance of behavior. This element is not essential for a "descriptive" definition of values, or what values are. It is of transcendent significance for a functional definition -- what values do. This meaning of value-as-criterion treats values as indirect determinants of all attitudes and actions, whether to adjust to one's physical and social environment, enhance and defend his self-esteem and social conscience, or improve knowledge and growth. It means that attitude objects that perform only limited service as criteria for conflict resolution and decision-making, or that affect only a narrow domain of life situations, would not be defined as values.

(8) A value is derived from, and modified through, personal, social, and cultural learning. Whereas the seventh element emphasized the consequences of a value, the final element underscores the antecedents of a value. It indicates how values originate and change. Being learned early in life, values are historically primitive, probably harder to change than the more specific attitudes and behavior patterns that are derived from them, and are a major source of stability both for the individual and society. However, once they are shifted themselves or severely deprived, they can also be sources of major upheavals. Values thus possess an extremely high degree of centrality in a person's life.

To summarize, any object of interest to a consumer may fall within the definition of a consumer value provided it meets the criteria mentioned above. It can involve any generalized person-centered end-state of existence, mode of behavior, pattern of ideas, grouping of people, biological condition, or direct benefit that is highly esteemed by the consumer himself, or perceived to be esteemed by others who are important to him. It is used by him as a criterion for evaluating a wide variety of other values, objects, activities, and people. It is an enduring type of belief, widely held by the consumer and by many but not all other members of his subculture or society. It is acquired and modified through personal, social, and cultural learning starting early in life, and plays a central role in his decision-making.

References

- C. E. Ayers, "The Values of Social Scientists," Journal of Social Issues, 6, (1950), pp. 17-18.
- H. Becker, "Supreme Values and the Sociologist," American Sociological Review, 6, (1941), pp. 155-172.
- S. L. Bem and D. J. Bem, "Training the Woman to Know Her Place: The Power of a Nonconscious Ideology," in H. H. Kassarian and T. S. Robertson, eds., Perspectives in Consumer Behavior, (Glenview, Illinois: Scott, Foresman and Company, 1971).
- P. Blau, Exchange and Power in Social Life, (New York: Wiley, 1964).
- C. J. Clawson, "Lewin's Vector Psychology and the Analysis of Motives in Marketing," in R. Cox and W. Alderson, eds., Theory In Marketing (Chicago: Irwin, 1950), pp. 41-63.
- C. J. Clawson, Psychology in Action (New York: Macmillan, 1946).
- M. T. Copeland, Principles of Merchandising (Chicago and New York: A. W. Shaw Company, 1924).
- E. Durkheim, The Division of Labor in Society (Glencoe, Illinois: Free Press, 1960).
- R. O. Evans, "Personal Values as Factors in Anti-Semitism," Journal of Abnormal and Social Psychology, 47 (October 1952), pp. 749-756.
- R. Frondizi, What Are Values: An Introduction To Axiology (LaSalle, Illinois: Open Court, 1963), p. 4.
- N. Georgescu-Roegen, "Utility," in E. Sills, ed. International Encyclopedia of the Social Sciences, Vol. 16, (New York: Macmillan, 1968), pp. 236-267.
- F. Hansen, "Psychological Theories of Consumer Choice," Journal of Consumer Research, 3(1976), pp. 117-142.
- L. H. Haney, History of Economic Thought, (New York: The Macmillan, 1936).
- R. S. Hartman, The Structure of Values: Foundations of Scientific Axiology (Carbondale, Illinois: Southern Illinois University Press, 1967).
- W. A. Henry, "Cultural Values Do Correlate With Consumer Behavior," Journal of Marketing Research, 13 (1976), pp. 121-127.
- J. A. Howard, Consumer Behavior: Application of Theory (New York: McGraw-Hill, 1977).
- L. M. Irelan and A. Besner, "Profile in Poverty," in H. H. Kassarian and T. S. Robertson, eds., Perspective In Consumer Behavior (Glenview, Illinois: Scott, Foresman and Company, 1971)
- J. Jacoby, "Consumer Psychology: An Octennium," in M. R. Rosenzweig and L. W. Porter, eds., Annual Review of Psychology, 27(Palo Alto: Annual Reviews, Inc., 1976), pp. 331-358.
- H. H. Kassarian and T. S. Robertson, eds., Perspective In Consumer Behavior (Glenview, Scott, Foresman and Company, 1971), p. 292.
- E. J. Kelley and W. Lazer, Managerial Marketing (Homewood, Illinois: Irwin, 1973), p. 113.
- Kurt Lewin, "Forces Behind Food Habits and Methods of

- Change," Bulletin of the National Research Council, 108 (1943) pp. 35-65.
- C. Kluckhohn, "Values and Value Orientations in the Theory of Action: An Explanation in Definition and Classification," in T. Parsons and E. Shils, eds., Toward A General Theory Of Action (Cambridge: Harvard University Press, 1951), p. 395.
- F. R. Kluckhohn and F. L. Strodtbeck, Variations In Value Orientations (Evanston: Row, Peterson and Company, 1961).
- D. Lee, "Culture and the Experience of Values," in A. H. Maslow, ed. New Knowledge In Human Values (New York: Harper and Row, 1959), p. 165.
- G. Lundberg, "Human Values: A Research Program," Research Studies (Washington State University), 18 (1950), pp. 103-111.
- A. H. Maslow, Motivation and Personality (New York: Harper & Row, 1954).
- R. Mayer, ed., Changing Marketing Systems: Consumer, Corporate and Government (Chicago: American Marketing Association, 1967).
- W. J. McGuire, "Psychological Motives and Communication Gratification," in J. G. Blumler and E. Katz, eds., The Uses of Mass Communications: Current Perspectives on Gratifications Research (Beverly Hills: Sage Publications, 1974).
- P. McReynolds, Four Early Works on Motivation (Gainesville, Florida: Scholars' Facsimiles and Reprints, 1969).
- A. Michell, "Changing Values and the Marketplace," in F. C. Alluine, ed., Relevance in Marketing: Problems, Research and Action (Chicago: American Marketing Association, 1971).
- J. M. Munson, "Personal Values: A Cross-cultural Assessment of Self Values and Values Attributed to a Distinct Cultural Stereotype," Forthcoming, Proceedings, Association for Consumer Research, 1977.
- H. A. Murray, Explorations in Personality: A Clinical and Experimental Study of Fifty Men of College Age (New York: Oxford University Press, 1938).
- T. Parsons, "Pattern Variables Revisited," American Sociological Review, 25(1960), pp. 467-483.
- T. Parsons, The Social System (Glencoe, Illinois: Free Press, 1951).
- S. H. Patterson, Readings in the History of Economic Thought (New York: McGraw-Hill, 1932).
- R. Pinter, "A Comparison of Interests, Abilities and Attitudes," Journal of Abnormal and Social Psychology, 27 (January-March 1933), pp. 351-357.
- J. P. Robinson and P. R. Shaver, eds., Measures of Social Psychological Attitudes (Ann Arbor: Institute for Social Research, 1971) pp. 402-405.
- Milton Rokeach, "A Theory of Organization and Change Within Value-Attitude Systems," Journal of Social Issues, 24(1968b), pp. 13-33.
- Milton Rokeach, Beliefs, Attitudes, and Values (San Francisco: Jossey Bass, 1968a).
- Milton Rokeach, The Nature of Human Values (New York: The Free Press, 1973).
- Milton Rokeach, "The Role of Values in Public Opinion Research," Public Opinion Quarterly, 32(1968-69).
- M. J. Rosenberg, "Cognitive Structure and Attitudinal Affect," Journal of Abnormal and Social Psychology, Vol. 53 (November 1956).
- B. Russell, What I Believe (London: K. Paul-Trench, Trubner & Company, 1925).
- "Second Washington Social Indicators Conference to Explore the Response of Business to New Values," The Marketing News, 6(1973), p. 1.
- W. A. Scott, "Cognitive Consistency, Response Reinforcement, and Attitude Change," Sociometry, 22 (1959), pp. 219-229.
- P. A. Sorokin, Society, Culture and Personality, Their Structure and Dynamics: A System of General Sociology (New York: Cooper, 1947).
- E. Spranger, Types of Men: The Psychology and Ethics of Personality-Translated from German by P. J. Pigors (Halle: Max Niemeyer Verlag, 1928).
- W. I. Thomas and F. Znaniecki, The Polish Peasant in Europe and America (New York: Alfred A. Knopf, 1927) p. 1131.
- L. L. Thurstone, "The Measurement of Values," Psychological Review, 6(1954) pp. 47-58.
- A. Toffler, "Value Impact Forecaster--A Profession of the Future," in K. Baier and N. Rescher, eds., Values and the Future (New York: The Free Press, 1969) p. 4.
- F. Tonnies, Community and Society, Gemeinschaft and Gesellschaft (East Lansing, Michigan: Michigan State University Press, 1957).
- R. Turner, "Modes of Social Ascent Through Education: Sponsored and Contest Mobility," in A. H. Halsey, J. Floud and C. A. Anderson, eds. Education, Economy and Society (New York: Collier-Macmillan, 1961).
- D. E. Vinson, "Personal Values as a Dimension of Consumer Discontent," in B. Greenberg and D. Bellenger, eds., Contemporary Marketing Thought (Chicago: American Marketing Association, 1977) p. 505.
- D. E. Vinson and J. M. Munson, "Personal Values: An Approach to Market Segmentation," in K. L. Bernhardt, Marketing: 1776-1976 and Beyond (Chicago: American Marketing Association, 1976), pp. 313-317.
- D. E. Vinson, J. M. Munson and M. Nakanishi, "An Investigation of the Rokeach Value Survey for Consumer Research Applications," Proceedings (Atlanta: Association for Consumer Research, 1976).
- D. E. Vinson, J. E. Scott and C. M. Lamont, "The Role of Personal Values in Marketing and Consumer Behavior," Journal of Marketing (1977), pp. 44-50.
- W. H. Werkmeister, Man and His Values (Lincoln: University of Nebraska Press, 1967).
- W. L. Wilkie and E. H. Pessemier, "Issues in Marketing's Use of Multi-Attribute Attitude Models," Journal of Marketing Research, 10 (1973), pp. 428-441.
- A. D. Woodruff and F. J. Divesta, "The Relationship Between Psychological Measurement," Psychological Bulletin 8 (1948), pp. 645-660.

VALUES AND CONSUMPTION PATTERNS: A CLOSED LOOP

James M. Carman, University of California, Berkeley

Abstract

The author had proposed a model of life style and consumer behavior that relies strongly on the concept and measurement of values as advanced by Milton Rokeach. Here the work is extended to make it a more comprehensive model of consumer decision making and to make explicit a set of feedback loops to values and life style. The model is explained and some examples of recently emerging life styles are developed.

Introduction

For over a decade, my work in the differentiation of life styles has had a stronger anthropological flavor to it than that of most other researchers in marketing. In the past two years, I have been attracted by the ability of Milton Rokeach, through his work in values, to provide a social psychological basis for my anthropological approach to this topic (Rokeach, 1973). In this paper I will attempt to extend this line of theoretical development by: (1) adding a bit more sociological insights to it and (2) showing how the link from values to subcultures to consumption patterns is actually a loop which feeds back to values, life styles, and subculture.

In this first section, my framework for integrating values, cultures, life styles, and consumption is summarized. In the second section, the model is extended. In the third section, examples of how subcultures emerge and then survive or decline and the importance of the feedback loop in this process will be explained. This section also suggests some public policy implications of this line of research.

My interest in this area stems from my early work in the investigation of social class as a basis for market segmentation. At that time I was critical of the sociological view of social class based largely on status groups or socioeconomic factors such as education. My position was that the groups of interest for the study of consumption patterns were really subcultures. The richness of this area for market segmentation lay in the area of identifying subcultures within the society that tend to be self-perpetuating and are homogeneous with regard to their values, beliefs, motivation, interests, activities, and thereby media exposure behavior. This early literature is summarized in Carman (1965) and Frank, Massy and Wind (1972).

Despite these findings, the social class reviews and research since 1968 largely ignore these conclusions. In some of this work, the construct is not defined; in others it is the idea of status class. It is little wonder that this work under the name of social class has not persisted.

Instead, work along these lines proceeded under such names as "activities, interests, and opinions," "life style," and even "psychographics." This literature has been summarized by Hustad and Pessemier (1974) and Wells (1975). Despite all this flurry of activity, there continues to be nagging questions as to whether there is any theoretical underpinnings to this life style research.

Thus, I was delighted when the work of Rokeach on values seemed to provide a social psychological basis

for this area as I conceptualized it. The framework I developed, shown in Figure 1, is based on a cause and effect relationship between four sets of variables (Carman, 1976/77). Terminal values lead to holding certain instrumental values to be more important than others. These are the causes of interests and assumed roles that, in turn, are the cause of the activities in which one uses time. Each set of variables--terminal values, instrumental values, interests and roles, time use activities--influence consumption behavior both directly and through the other intervening variables. Consumption behavior is defined here to include purchase patterns, search behavior, and media exposure behavior.

More recently, I have added feedback links to indicate that some purchases, search, and media behaviors have an effect that feeds back to alter some values, roles, interests, and activities. It is this area that my thinking of a year ago now appears overly simplistic. The heart of this paper is a strengthening of how the sociological and anthropological inputs to the society also must be at work to determine just which configurations or constellations of values will exist in great enough numbers for a long enough period to be of interest to consumer behaviorists. But first, I believe it would be useful to review just a bit more the problems and experiences to date in measuring values with an emphasis on validity of the Rokeach instrument.

The Measurement of Values

I am struck with the frequency with which we hear and read statements to the effect that little systematic work has been done on the measurement of values. Anthropologists have clearly been interested in linking the study of cultures with the values held by the people of a culture, but anthropologists, working as they do primarily after the fact, are not likely to develop instruments that will satisfy other social scientists. In both sociology and social psychology, values have remained a rather isolated area of study. The reviews of work prior to Rokeach do not build confidence (Dukes, 1955; Robinson and Shaver, 1969). In other fields, the word values is used in unfortunate ways that could be the source of much confusion. In organizational behavior, for example, the word values is used to connote attributes of a job, e.g., interesting, relative wage, advancement opportunity (Pennings, 1970).

However, it appears to me that much of this confusion and criticism can not be applied to Rokeach. Robinson and Shaver state that their review of the literature shows that all values fall into five broad categories: (1) means vs. ends, (2) good vs. evil, (3) aesthetics, (4) intellectual, and (5) economic. Rokeach includes all of these categories in his instrument. Similarly, the Bales and Couch Value Profile (1969) that received some validity from the work of Withey (1965) includes dimensions that also are tapped by Rokeach. At these meetings last year, Vinson, Munson, and Nakanishi (1976) reported on a study that confirms the Rokeach instrument as tapping values that can differentiate among consumer groups. Also in marketing, Scott and Lamont (1973) have found the Rokeach framework useful.

Finally, later in this paper we will refer to a theory of Etzioni (1972) suggesting that when groups of

persons find their values not in harmony with the balance of their culture, they are likely to respond with changes of instrumental values, interests, activities, consumption, and media habits that follow patterns of: (1) hedonism, (2) substitution of symbols for things as relevant objects of value, (3) empathy and community with those of similar value incoherence, and (4) collective activist behavior. While Etzioni did not do it, his theory is completely testable with the Rokeach instrument. Thus, it is not true that the Rokeach Value Survey is just another attempt to measure values that is on thin ice. It has had an incredible amount of testing in a wide variety of applications. One may want to modify the scales in particular applications, but we believe we can convince you that Rokeach provides a framework and a measure of values that are applicable in anthropology, sociology, psychology, and consumer research.

Extension of the Model

I said earlier that the addition of simple feedback loops to Figure 1 oversimplifies the link between values, life styles, and consumption. Certainly models of consumer segmentation like that of Wilkie and Cohen (1977) that link demographic variables, socioeconomic variables, and personality variables as causes of media exposure, life style, and AIO are oversimplified to the point that I believe they are not useful for theory development. To demonstrate this point, we will enrich our model by the inclusion of more sociological concepts.

First, it will be useful to be more specific about the definition of life style. The term is not used in Figure 1. Nor, for that matter, is "subculture." A subculture is a collective which shares values. It has considerable consensus with respect to values, meanings, and the role of social institutions. It is enduring and self-perpetuating. It may exist across a range of socioeconomic classes, but it is a self-conscious collective. Its members consciously identify with it. Southern white Anglo-Saxon Protestants are an example of a subculture.

Although the term is poorly defined in the contemporary literature, life style is probably best thought of in terms of activities, interests, and opinions that become reflected in terms of media and consumption patterns (Zablocki and Kanter, 1976, 270-71). Figure 1 suggests that these patterns come from shared values. In marketing, we are interested in identifying segments of people who share life styles. By this definition, a subculture consciously shares a small number of distinct life styles.

The difference between a subculture and a life style is its enduring capabilities. A life style segment does not necessarily build up an institutional structure for endurance and perpetuation. It is identified by AIO and consumption behavior. Marketers are interested in identifying life style segments and constellations of values held by these segments. They and all social scientists are interested in tracing the dynamic development of a life style to see if it shows signs of developing into a new subculture. The transition of a life style will give clues as to how it might evolve and whether or not it will develop into a new subculture.

An example would be the "whole earth" movement--the return of young, healthy adults to a rural society that places an emphasis on living harmoniously with nature and on self-sufficiency. This is a life style and it is a viable market segment for many sellers. If you doubt this, I suggest you count the percentage of

advertising in the magazine, Whole Earth News. Will it endure? On the one hand, it appears to resist institutionalization. It is almost a life style of the hermit. On the other hand, there are signs of institutional development.

The point of this digression into life styles and subcultures is to point out that we must be careful we understand where values fit into a broader social structure if we are to find values useful in consumer research. Note that it is only a first step to show that there is a correlation between holding certain values and particular consumption patterns. Such a correlation may say something about the validity of the Rokeach Value Survey, but it does not take us very far toward a theory that relates values to consumption patterns. We need to do more work in three other areas to make headway in theory development. First, we need to build and test a theory that is richer in its specification of the linkage between values, other socioeconomic and demographic variables, cultural variables, life styles, and consumption. Second, we need to do more to identify general life style segments. Third, we need to have better explanations of the evolution of these life style segments. Figure 2 offers a small contribution toward these ends.

Figure 2 is an expansion of Figure 1, but it also contains some changes. These will be described first. The concept of life style now appears explicitly and is defined as roles, interests, time use activities, consumption, shopping, and media patterns. The life style construct is then treated as a single construct. Values continue to be principal inputs to life style. Three new factors have been added on the left side of the figure: subcultures, social institutions, and exogenous forces.

Subculture is shown as it was defined above, i.e., an enduring collective identified by its life styles and by a coherent value system perpetuated through an institutional framework. The Southern WASP is an example of a persistent subculture that has a unique constellation of values including conservatism, tradition loving, family loyalty, and religion (Reed, 1972). Reed hypothesizes that the institutions and individual values required to sustain this value constellation continue because southerners see themselves as a suppressed minority who need to maintain defense mechanisms against continued external pressures. Suppression as a factor encouraging endurance is not represented in Figure 2.

In contrast, social institutions appear to be of sufficient importance in the formation and perpetuation of value constellations that this factor has been added to the diagram. The representation is designed to show the importance of institutions as necessary reinforcers of specific values or constellations of values. Institutions form around or accept a particular value structure. They help the subculture they serve to maintain a consensus of coherent values. They act as orderly change agents to adjust value constellations through time. It should be emphasized that these institutions function well in accomplishing slow, orderly change. They are not change agents and respond badly to sudden changes thrust upon them. Established churches are the prime example of such institutions.

In order to simplify the feedback links in the diagram, the feedback to values is shown only to terminal values. Likewise, the feedback loop from consumption and media patterns is shown only to instrumental values. One should not reject the hypothesis that both feedbacks go to both kinds of values. Similarly, the feedbacks from purchases to values and life styles are not explicitly hypothesized.

FIGURE 1
THE ORIGINAL MODEL

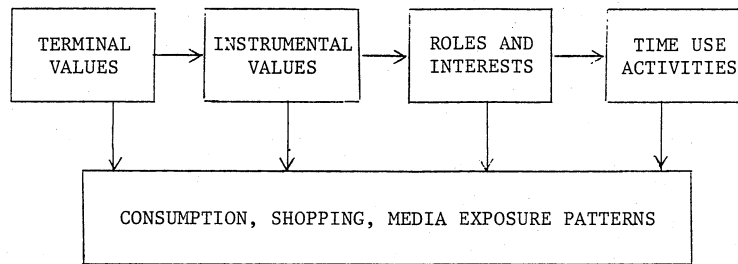
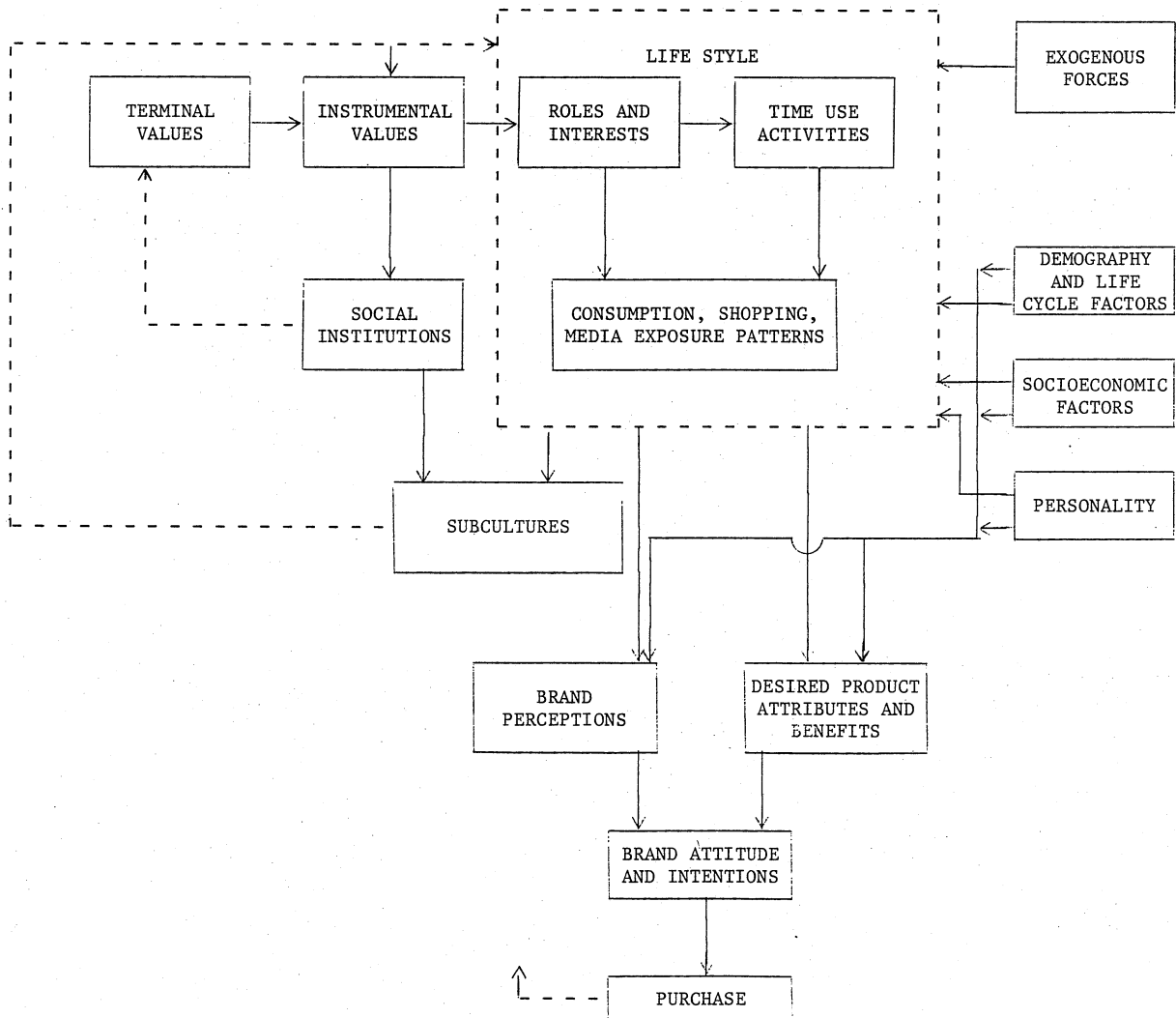


FIGURE 2
AN EXPANDED MODEL OF VALUES, LIFE STYLES, AND CONSUMPTION



The exogenous forces box has been added to indicate that often life styles can emerge from factors outside personal factors such as values, personality, life cycle, and socioeconomic factors. It is in this area that the anthropologist and sociologist have the most to contribute to the processes under study. Many anthropologists believe that new life styles proliferate during periods of societal breakdown (Kroeber, 1957). Thus, some hypothesize that the war in Southeast Asia in the 1960s was a major contributor to the emergence of unusual life styles.

Other exogenous factors that might be influencing life style development today are evident but not as often discussed. Increases in real income and the substitution of capital for labor is an interesting current example. A reasonable hypothesis for the emergence of a number of life styles that are more self-oriented and leisure oriented is that increased real income and leisure time have permitted life styles that previously were not possible (Gartner and Riessman, 1974; Nicosia, 1974). If this were true, then one would expect to observe an increase in work-oriented values and life styles during recession and an increase in play-oriented life styles during prosperity. While not formally tested to my knowledge, this is a plausible testable hypothesis useful to consumer research and greatly aided by the measurement of values.

For example, consider social programs designed to urge people to conserve energy and, in the West, water. Is there value congruence between the program and existing or emerging values and life styles? In California it appears that water conservation is more successful than gasoline conservation. People with different value constellations may have different responses to conservation efforts between the two commodities. Were this to be true, then the appeals used to achieve conservation should be different and should be based on an analysis of differences in values.

Returning to the description of Figure 2, look next at the three factors on the right below "exogenous factors." These three, demographic factors, socioeconomic factors, and personality, are the conventional determinants of purchase behavior one finds in the literature. They are added here for two reasons. First, for comprehensiveness. For a complete understanding of the influence of values and life style on consumption, it is necessary to integrate these concepts into a comprehensive model. Second, it is important to see that it is suggested that these three conventional determinants influence consumer behavior in two ways. One is the conventional direct links to brand perceptions and to desired benefits and attributes. I call this conventional for it appears this way in the popular comprehensive models.

The second way is through life style. Some life styles we observe are highly dependent on life cycle. For example, there are life styles that attract only the young and others that appeal to only the retired. There are other life styles that are dominated by economic considerations. For example, there is increasing concern in this country for the "underclass," that tenacious subculture on long-term welfare who seem to be drifting farther and farther from the mainstream of American life. Finally, there is a long tradition in sociology of research in occupational impact on life style (Kanter, 1976b).

This double effect of the conventional determinants of consumer behavior--directly and through life style--is not original with this model. It has been suggested by others. Plummer (1977), for example, has suggested a similar idea based on a theory of George Kelly (1955) that people make decisions using two construct systems:

"superordinate constructs" are those stable constructs that flow from life style and subculture and "subordinate constructs" that are short-term, problem-solving constructs.

Before closing this description of Figure 2, it may be useful to make one comment about the lower boxes on brand perceptions, brand attributes, attitudes, intentions, and purchase. It was not our desire here to be innovative or controversial. It did, however, seem desirable to tie our value and life style discussion into conventional models of individual choice decisions.

Closing the Loop

In this last section, the model will be exercised a bit by application of Etzioni's (1972) four life style responses to value incongruence. Earlier in this paper these responses were introduced in terms of values. The four responses are renamed using Rokeach terms rather than Etzioni terms. They are an increase in the importance of the values: (1) hedonism--a pleasurable, enjoyable life, happiness and contentment; (2) etherealization--inner harmony, a world of beauty, culture and art; (3) community--true friendship, equality among men; (4) activism--a lasting contribution, social consciousness. (Some of these terms are those of Zablocki and Kanter, 1976.) In each case, it is possible to trace a change in value constellation to a new life style to a consumption pattern and potentially a feedback to new values. In addition, each raises some interesting public policy questions that would yield to research on particular areas of values and life styles. This exercise may provide the reader with a bit more understanding of how the model operates and some of the research topics that flow from the model.

Hedonism

Etzioni (1972, 8) states that hedonism develops when the normative values of a society lose meaning and disintegrate without being replaced by new norms. What life styles and consumption patterns are associated with an increase in popularity of this value constellation? Sexual freedom as an ideological position, the drug culture, self-fulfillment, encounter groups, EST, human potential movements, and a personal shrink are all consumption efforts aimed at replacing societal norms with an assurance of being right with one's self. A better understanding of these trends in life styles could lead to a better understanding of rational public policy toward community mental health and drug abuse programs.

Etherealization

Etherealization is the substitution of symbols for things as objects of value. This shift follows disenchantment with the value of capital accumulation and with an unwillingness or inability to work hard enough to acquire "things." Symbols are relatively costless and in infinite supply. The practice has often been observed among the very poor, for example, in Latin America. The life style and consumption patterns that emerge from this new value constellation are an increase in the popularity of fundamentalist religions, mystical Eastern religions, religions based on poverty, natural foods, ecology and natural beauty. Note that the increase in aesthetic values has led to considerable pressures on government for greater support of open spaces, endangered animals, arts, and continuing education. Understanding the nature of these changes in values should lead to a more rational method for prioritizing government programs in these areas. Note that these new religions appeal to the former constellation (via the inner harmony route) and to the next

constellation (via the monastic route).

Community

People may seek to replace old value norms with new ones by clustering with other persons perceived to share similar value constellations. These might be existing communities, i.e., one might return to a known institution or community when the outside world looks frightening or one might seek to find a new utopian community. This decade has provided us with many examples of both kinds, but certainly closed communal enclaves with some ethnic and religious homogeneity have experienced periods of popularity at many times in the history of the western world. Clearly, communal living results in drastic life style change. Community need not always take the form of a commune or be in a rural setting. Gangs of youths in low-income inner-city areas also fall into this category. They too are not unique to the sixties and seventies, and also result from value incongruence. Note that forming physical community is a technique for institutionalization that could lead to a subculture with a normative set of values and an institutional structure for reinforcing these values. A better understanding of urban community life styles could lead to an improved public policy toward a whole host of poverty problems including frustration, personal bankruptcies, crime, and the disenfranchised consumer.

Activism

When the formation of a physical community is not possible, one alternative is a community organized for political activism. Thus, again, these four responses are not mutually exclusive. Many communities are politically active and many activists work through communities. A political or social movement is not a life style, "but because a social movement involves an ideology that orients its followers toward a coherent set of values, the context of the social movement generally provides an opportunity for a life-style to develop around it" (Zablocki and Kanter, 1976, 292). Again, the movement might become institutionalized to the point that it is value reinforcing.

References

R. Bales and A. Couch, "The Value Profile: A Factor Analytic Study of Value Statements," Sociological Inquiry, 39(1969), 3-17.

James M. Carman, The Application of Social Class in Market Segmentation (Berkeley: Institute of Business and Economic Research, University of California, 1965).

James M. Carman, "Life Style Segmentation: Is There a There There?," Australian Marketing Researcher, 1 (Summer, 1976/77), 53-62.

W. F. Dukes, "Psychological Studies of Values," Psychological Bulletin, 52(1955), 24-50.

A. Etzioni, "The Search for Political Meaning," Center Magazine, 5(March/April, 1972), 2-8.

Ronald E. Frank, William F. Massy, and Yoram Wind, Market Segmentation (Englewood Cliffs: Prentice-Hall, 1972).

A. Gartner and F. Riessman, The Service Society and the Consumer Vanguard (New York: Harper & Row, 1974).

Thomas P. Hustad and Edgar A. Pessemier, "The Development and Application of Psychographic, Life Style and

Associated Activity and Attitude Measures," in W. D. Wells (ed.), Life Style and Psychographics (Chicago: American Marketing Association, 1974), 31-70.

R. M. Kanter, Work and Family in America (New York: Russell Sage Foundation, 1976).

George A. Kelly, The Psychology of Personal Constructs (New York: W. W. Norton, 1955).

F. N. Nicosia, Summary Report. Technology and Consumers: Individual and Social Choices, Report to the National Science Foundation (Berkeley: Institute of Business and Economic Research, University of California, 1974), 61-63.

J. M. Pennings, "Work-Value Systems of White-Collar Workers," Administrative Science Quarterly, 15(1970), 397-405.

Joseph T. Plummer, "Life Style, Social, and Economic Trends Influencing Consumer Satisfaction," in H. K. Hunt (ed.), Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction (Cambridge: Marketing Science Institute, 1977), 382-408.

John S. Reed, The Enduring South: Subcultural Persistence in Mass Society (Lexington, Mass.: Heath, 1972).

John P. Robinson and Phillip R. Shaver, Measures of Social Psychological Attitudes (Ann Arbor: Institute for Social Research, 1969).

Milton Rokeach, The Nature of Human Values (New York: The Free Press, 1973).

Jerome E. Scott and L. M. Lamont, "Relating Consumer Values to Consumer Behavior: A Model and Method for Investigation," in Thomas V. Greer (ed.), Increasing Marketing Productivity (Chicago: American Marketing Association, 1974), 283-88.

Donald E. Vinson, J. M. Munson, and Masao Nakanishi, "An Investigation of the Rokeach Value Survey for Consumer Research Applications," Proceedings of the 1976 Conference (Chicago: Association for Consumer Research, 1976), 247-52.

William D. Wells, "Psychographics: A Critical Review," Journal of Marketing Research, 12(May, 1975), 196-213.

William L. Wilkie and Joel B. Cohen, "An Overview of Market Segmentation: Behavioral Concepts and Research Approaches," Report No. 77-105 (Cambridge: Marketing Science Institute, 1977).

S. Withey, "The U.S. and the U.S.S.R.: A Report of the Public's Perspective on the United States--Russian Relations in Late 1961," in D. Bobrow (ed.), Components of Defense Policy (Chicago: Rand McNally, 1965), 164-74.

Benjamin D. Zablocki and Rosabeth M. Kanter, "The Differentiation of Life-Styles," in A. Inkeles, J. Coleman, and N. Smelser (eds.), Annual Review of Sociology (Palo Alto: Annual Reviews, Inc., 1976), 269-298.

THE EUROPEANIZING OF AMERICA:
A STUDY IN CHANGING VALUES

Donald L. Kanter, University of Southern California

Abstract

The theses is that technological, environmental, governmental and economic changes in America have produced changes in terminal and instrumental consumer values. Metaphorically speaking, these changes appear to approximate the European mode in terms of their social-psychological implications. These specific value changes are identified along with some of their marketing and advertising consequences.

Introduction

With American products, fast food franchises, and United States multinational corporations girdling the globe, one has customarily thought in terms of the Americanizing of Europe, not the opposite. Changes, however, in the external conditions in America seem to have occurred producing changes in terminal and instrumental consumer values and attitudes (Rokeach, 1973; Rokeach, 1960).

Interestingly enough, this rapid shift in socio-psychological outlook seems to more or less approximate a European mode in terms of behavior as well as weltanschauung. It should not be startling to say these changes have vast implications for consumer marketing strategy (Katona, 1971).

There will be fewer retreats for Americans to former days of conspicuous consumption. External conditions--technological, psychological, environmental and governmental--are the constraints.

For instance, government has become an anonymous, regulatory force in American life with all the depersonalization this implies. As the bureaucratic structure, described by Max Weber and Prof. Merton, has mushroomed, Americans have come to question the immediacy and responsiveness of the Democratic process. Recent candidates of both major political parties have campaigned on this point.

Moreover, Americans no longer perceive that they control the world, so to speak. European countries lost their colonies long ago and felt the destruction of two world wars. It is only recently, however, that Americans have felt diminished power over world affairs, with all of the concomitant threats to security and potency this implies. Terrorism, along with diminished control, has also become part of the collective unconscious, adding to feelings of general slippage and unrest.

To add to this decline of perceived omnipotence is the realization and acceptance of finite and limited resources. It is, for example, a significant development that American oil companies have undertaken massive advertising campaigns urging consumers to conserve their product, petroleum, not to consume it.

Under these changing conditions, marketing premises posited over the last twenty years of the "Marketing Revolution" are virtually invalid. Those assumptions of unlimited affluence, boundless food, fuel, and productivity; and--most importantly--a vision of the American consumer as voracious, eager, hungry, and greedy, no longer seem accurate (Webster, 1974).

The future has arrived bringing changed aspirations and new social and sexual roles. Americans will make new attempts to cope with increasing depersonalization and anonymity brought about by over-population and bureaucracy as well as inflation and psychological diminution. People are seeking different means of achieving equilibrium in a continuously clamorous world, bringing about different ways of consuming and evaluating goods and services.

It is in these differences--the response to a changing total environment--that the Europeanizing of America may be seen. Clearly, the analogy is imperfect; but it may be descriptive of a new values constellation. As conditions of scarcity, bureaucracy and inflation--recurring European experiences--occur in this country, the modes of adjustment appear to become closer in substance and style.

The Changes

First there is the American loss of innocence. Something very simple is meant by this loss of innocence which goes beyond Mark Twain, Edith Wharton or Henry James. Americans never before had military and diplomatic defeats approaching the magnitude of Viet Nam. They had not recently come close to a Watergate or an international oil cartel, or known crime and violence to such an unprecedented and televised extent.

What is the significance, however, of loss of innocence to the Europeanized American? It means, among other things, that what is vulnerable and uncertain about life has been brought home to most Americans. Certain mainstream cultural assumptions lost validity in the late 60's and the life of everyday quiet desperation is not so easily tolerated among many people. People feel they cannot manipulate their destinies as much, formerly an article of faith. This contributes to feelings of anomie. Psychologists might say that the locus of control has shifted from the self to external forces (Thornbill, 1975).

America's adult generation was raised to believe that one can be virtually anything he wants to be, including President of the United States. But Americans are moving towards a more fatalistic attitude, resembling that found in Europe ("que sera, sera"). There is a lowered expectation about the ability to manipulate one's environment. In addition, the parameters have been narrowed in regard to the invincibility of self-help. Boundless optimism in achieving traditional American goals seems inappropriate.

A related aspect of Europeanizing is the enormous growth in skepticism and cynicism. People question other people's motives more than before. The concept of the blunt, open-faced American, a gulleless, Candidate-like character, is not a valued aspiration. In the current environment this does not help survival.

What does this increased cynicism mean? For one thing, authority figures are suspect. A study conducted at the University of Southern California to examine reactions to public service advertisements about illegal drugs is illustrative (Kanter, 1974).

Star athletes (ostensibly reference group symbols) in

effect said: "I can't do my thing if I'm high on drugs." These ads elicited disbelief in 88% of the grade school children tested. When questioned further, the students felt one can't play that kind of game without being high and that the revered athletes were lying.

The current American approach to skepticism and cynicism may be characterized in part as trading off. This approach recognizes that things are rarely absolute, ("What do I have to give for what I get?") People appear to have arrived at a certain shrewdness in their dealings with this realistic and cynical approach to life. "Street smarts" have become operative in our society. In a phrase, this is the ability to "aim off"--to take one's time. Don't quite reveal oneself; maintain a facade; feign ignorance. This is the time-honored peasant's protection.

Another basic characteristic of Europeanizing is the changing time-frame of Americans. No longer is immediate response a virtue. From the customary American spontaneity and rapid closure, people tend to feel that it is notoriously unsafe to enter into a situation, to make decisions without waiting for events to clarify. This is an old English technique of negotiation identified by Charles Dickens in Bleak House in the 1860's, when he spoke of "fogging the issue." In other words, conditions change, hence wait, it may even go away--if not, continue to obscure the point and postpone a decision.

But a most crucial element of this phenomenon is that people want time--almost more than things. The marketing implications of this are enormous. Time will be a reward assuming many forms--longer holidays and shorter workweeks; temporal instead of financial bonuses.

The Europeanizing of Americans is also expressed in new levels of aspiration. Not everyone wants to grow up to be President or leader. Many would prefer simply to "play the game," collect time and have fun, being cooler and less hassled. Increasingly, aspirations are not so much power oriented as time oriented: Narcissism or self-absorption is the new patriotism, up to a point. This is absolutely crucial to the leisure, automotive, cosmetic, and a multitude of other businesses.

The trade-offs between risk taking and security are also changing. As corporations--including multinationals--grow, loyalty does not tend to keep pace, it is actually quite limited. People prefer to keep their options open. By "hanging loose," to use the argot, one doesn't have to commit all that much of oneself. One may try something, stay detached, then move--a shrewd "street-smart's" way of operating. Literally, take the money and run.

Also, there is a growing appreciation of natural simplicity and a European-type sensuality. Americans have come to realize the existence of basic body truths; that biological man is a very natural entity, smells and sights included.

In terms of sex, for instance, more people wallow less in guilt than in pleasure, no matter how transient or shallow. But discretion, rather than confrontation (I've sinned) is the style of coping with an extracurricular relationship because people accept that simple bodily functions are natural, not perverse.

Along with these simplicities come things like walking out, European promenading. Even in those American cities where inhabitants fear there's no place to walk safely after dark, the urge is transformed as it were into getting away; the Friday afternoon get-away is becoming a mass ritual. People are more pastoral rather than less. Camping is enjoying phenomenal growth in

popularity. Clothing for the rugged outdoor life--blue jeans, hiking boots, work shirts--is high fashion.

The natural and simplistic trends have led, among other things, to a reevaluation of the foods one eats. The empty calories of junk foods are now part of the mental geography of most people, not only mothers. "Health" foods, as they were called, are now everyday fare.

Notions of quality and obsolescence have also undergone changes re-emerging in highly European form. More and more, Americans purchase goods for reasons of durability and repairability not just for the sake of change, even if they pay more initially. Automobiles are expected to last longer; refrigerators are expected to last longer. Things are nurtured not discarded.

This emerging mentality is an expansion of the idea that less is more. Things and their accumulation bring far fewer rewards than do accumulation of time, leisure, and freedom in being one's self.

Related to simplicity and the downplay of accumulation is a new esteem for physical fitness. Exhilaration for a bike ride or a ski slope are just as important as shopping (with all its psychic transactional satisfactions) for the odd piece of furniture. This means, among other things, new channels for dollars and opportunities to express individuality in even eccentric kinds of ways.

Europeans have lived with scarcity, limited resources and minimal locus-of-control feelings and have developed other means for asserting their dignity and individuality. This is the situation Americans are not dealing with. As Americans come to face limits on their positions, possessions and power, their quest for viable modes to assert themselves becomes crucial. This phenomenon of self-assertion under new environmental conditions can impact the advertising and marketing processes. This question will not be examined using Advertising as an example.

Advertising Strategy and the New Values

How may these insights about the emerging state of the consumer mind be used to further communications between manufacturer and consumer? The position this paper takes is that understanding the values orientation of the consumer should help to develop more sympathetic and instrumental approaches to communication from manufacturers which are perceived as such by consumers (Delano, 1975).

The opportunity lies in the recognition that people are actively searching for an identity in an increasingly depersonalized world. In this process, they're trying to discover not only themselves, but also these goods and services which might give them a certain individuality and uniqueness. It is as if the Americans are using goods and services to reach a sharper level of separation from their homogenized peers. In a comparatively materialistic society, the substitution of "things" for spirituality is probably a natural if not altogether salutary outgrowth of on-going values trends.

Helping consumers to discover elements of a brand's uniqueness (to them) is really a major function of advertising and communications. Advertising drops cues, hints and generally leads the consumer to rethink his relationship to brands that he may have taken for granted.

By rethinking his relationship he is, in effect, projecting an element of discovery of "me-ness" upon a brand product or service. In many cases, it is not enough (or too much) to tell people directly what brand benefits are: It probably is a more reinforcing learning process for the consumer to discover for himself

that there is a fit between a brand and himself. Discovery for oneself becomes critical when the perception of authority figures is unfavorable as it is in the present values configuration.

In other words, to communicate effectively within the context of current values the consumer must intrude himself, aided and guided by advertising cues and suggestions. This Discovery Process is one way out of anonymity and depersonalization into identity partly through goods and services. And therein lies its enormous attraction to the consumers of the 1980's who do value uniqueness even though they are pessimistic about achieving it.

Rightly or wrongly, then, consumption is one of the last remaining avenues of self-expression in most Western cultures. To quote Lester A. Delano, President of Campbell-Ewald International:

"Viewed this way, it seems evident that advertising's role is to tap the consumer's Discovery Process by continually surfacing new cues to help the consumer in his on-going search and to get him constantly to rethink his relationship to the brand instead of taking it for granted. There is ample evidence--ignored by many advertisers--that repeating the same cue (advertising) over and over is wasteful in the extreme."

"We need many inputs, many cues . . . but of the kind that helps people to discover things. The kind that gets them to project themselves into the message. To involve themselves in our selling argument, so to speak, we must personalize the message so that the consumer can respond personally and we must be unafraid to be slightly fuzzy, unclear, funny, incongruous or ambiguous . . . just as people are. In this way we can draw the consumer in, let him make his own closure."

The use of The Discovery Process techniques, e.g. ambiguity, novelty, incongruity, etc., has attraction for the consumers because it is a way through an increasingly computerized, depersonalized existence. The very process of searching for these goods and services while distinguishing one from the others is rewarding per se. The excessive preoccupation with self has gone so far as to be labeled by New York Magazine when they call it the "Me Decade." It is as if people are urgently trying to hivel themselves off from mediocrity, ordinariness and conformity.

Consumer Research and The Discovery Process

If one postulates that the thrust of marketing advertising productivity is to trigger The Discovery Process, then it is possible to measure advertising to see if in fact it does get people to rethink, reopen, or underscore their relationship toward the brand. In other words, it is possible to see if an advertisement--in pretesting conditions--does make for links to the self or the shock of recognition (Kanter, 1970).

It is possible to set up codes to classify response to advertising pretests which do in fact measure whether The Discovery Process has been activated. These codes, in rough terms, relate to at least the following:

- a brand's character in terms of expanding the consumer's fantasy
- personal problem solving
- projecting and elaborating the advertisement
- uncovering unsuspected relationships to the brand

- the rehearsal for purchase
- the embracing of symbology
- and the general process of dialogue with the advertisement.

The point is that each of the above responses may be measured and evaluated in the pretesting situation. This fact allows the manufacturer to be able to predict what will be sensible, appealing, and meaningful to his customers before spending a great deal of money on an actual campaign.

Conclusion

It has been the contention of this paper that significant changes in values are taking place in the United States which more often than not approximate the European model. Some of these tendencies have been described and some of their implications drawn. Moreover, the use of the techniques of The Discovery Process have been discussed as a way through general consumer depersonalization.

Finally, the measurement implications have been mentioned and the opportunity for pretesting developed.

It is indeed the aware and sensitive marketing person who understands his audience well enough to stay in appropriate touch with it. As values change in society so must the marketing approach.

References

- Milton Rokeach, The Nature of Human Values (New York: Free Press, 1973).
- Milton Rokeach, "A Theory of Organization and Change Within Value Attitude Systems," Journal of Social Issues, 24(1960).
- George Katona, Aspirations and Affluence - A Comparative Study (New York: McGraw Hill, 1971).
- Frederick E. Webster, Jr., Social Aspects of Marketing (New Jersey: Prentice-Hall, 1974).
- The Yankelovich Monitor (New York: Yankelovich, Skelly and White, Inc.).
- Michael A. Thornhill, "A Computerized and Categorized Bibliography on Locus of Control," Psychological Reports, (April, 1975), pp. 505-506.
- Donald L. Kanter, The Government Anti-Drug Campaign, A paper read to the American Psychological Association, (Honolulu, Hawaii, September, 1974).
- Lester A. Delano and Donald L. Kanter, "Over-Advertising: A creative Solution," Advertising Quarterly, (London, England, Spring 1975).
- Lester A. Delano, "Beyond DAGMAR," published A. N. A. Information Service, (New York, December, 1975).
- Lester A. Delano, "Fresh Ad Input Termed Key to Keeping Brand Alive," Advertising Age, (November 22, 1976), p. 70.
- Donald L. Kanter, "Communications Theory and Advertising Decisions," Journal of Advertising Research, (December, 1970).

CONSUMERS' REACTIONS TO SEX IN TV COMMERCIALS

Deborah K. Johnson, Needham, Harper & Steers
 Kay Satow, Needham, Harper & Steers

Abstract

This paper explores viewers reactions to sexual material in TV commercials. Findings from a large scale quantitative survey and a small scale qualitative study are reported and implications for advertisers are discussed.

Sex in Commercials

Last year's concern about sex and violence on television has switched to this year's concern about sex on TV. Newspapers and magazines are filled with articles about the new sexual permissiveness. But are consumers aware of this trend? Are they bothered by sex on the airwaves? And closer to our hearts, as an advertising agency, do they feel there is too much sex in commercials?

To answer these questions and more we fielded two studies: a quantitative lifestyle study and a qualitative follow-up to determine consumers reactions. Before presenting our results, we invite the reader to test his sexual I.Q. with the following test.

Sex Test

Please indicate whether you think each of the following statements is true or false. DO NOT copy from your neighbors.

- T F 1. Most people feel there is too little sex on commercials.
- T F 2. Older men are less concerned with sex on TV than younger men; thus the phrase "dirty old man."
- T F 3. Women who are concerned about sex on TV point to soap operas as prime offenders.
- T F 4. Unlike the discriminating younger women, older women hate all sexy commercials.
- T F 5. Many people don't find sexual innuendo offensive. They don't even know what it is.
- T F 6. Women expect sex in men's ads. After all men are beasts.

Do People Feel That TV Commercials Place Too Much Emphasis on Sex?

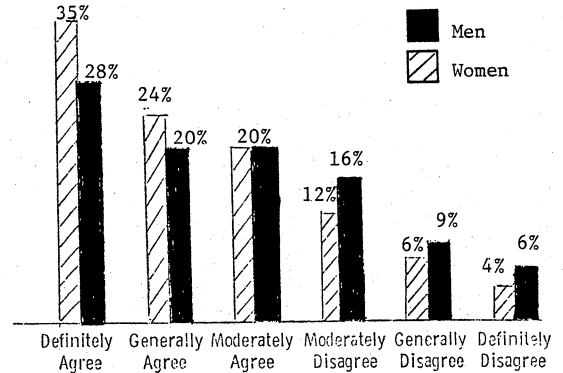
While published studies (e.g. TV Guide 1975) indicate that the majority of viewers object to the amount of sex on TV, we were not certain whether these objections extended to commercials as well as programs. For all we knew they may have felt there wasn't enough sex on commercials. To answer this question we turned to data from the Needham, Harper & Steers 1977 Life Style Study.

Every year NH&S mails a large scale consumer survey to 2000 married men and 2000 married women from the Market Facts' mail panel. Our response rate is high, averaging around 85% for women and 75% for men and the obtained sample matches the U.S. married population on age, edu-

cation, income and area of residence. The Life Style Study includes over 850 questions covering a wide range of activities, interests, opinions, product usage and media habits. Luckily for us one of the questions directly concerned attitudes toward sex on commercials. Respondents were asked to indicate on a 6-point scale the extent to which they agreed/disagreed with the statement, "TV commercials place too much emphasis on sex."

The responses to this question, shown in Table 1 clearly illustrate that the majority of our respondents agreed that TV commercials place too much emphasis on sex, and also that women tend to feel more strongly about this than men (79% of the women and 68% of the men showed at least moderate agreement with the statement).

TABLE 1
 "TV commercials place too Much Emphasis on Sex"



We discovered also that Age and Education are related to the degree of concern expressed over sex in commercials. As shown in Tables 2 and 3 strong agreement with the statement that "TV commercials place too much emphasis on sex" increases with age and decreases with education for both men and women.

TABLE 2
 % of Each Age Group That Definitely Agree That "TV Commercials Place too Much Emphasis on Sex"

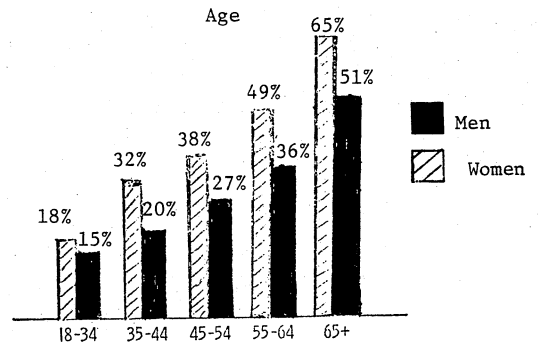
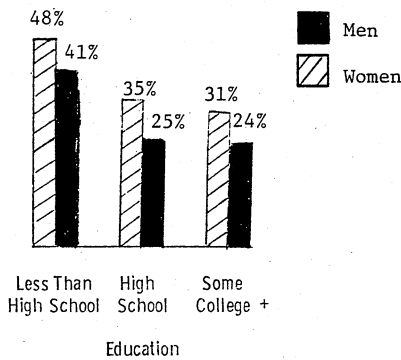


TABLE 3

% of Each Education Group that Definitely Agree That "TV Commercials Place too Much Emphasis on Sex"



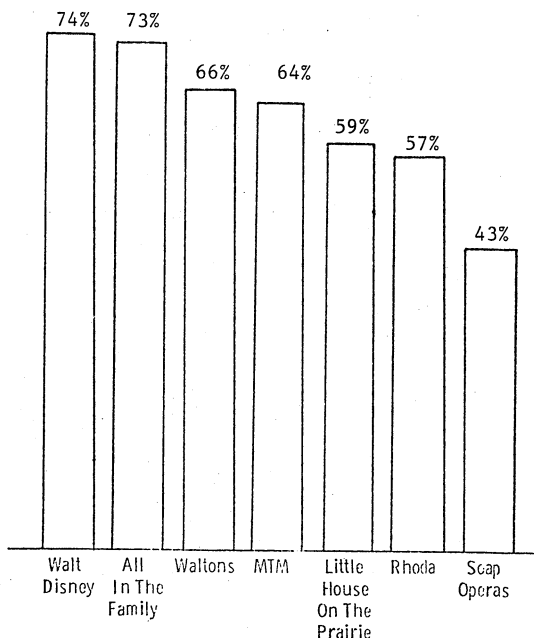
So far our findings made a good deal of sense and had presented us with no great surprises.

What TV Shows do People Who Are Bothered by Sex Watch?

The NH&S Life Style Study also asks about TV program viewership. We expected to find that TV preferences would reflect concern over the amount of sexual material on TV. For example, we expected concerned women to watch family shows, certain situation comedies and other wholesome TV fare. Table 4 presents some of the favorite TV programs of these women. Some of these shows made perfect sense to us. The Waltons, Walt Disney, Little House on the Prairie are all about as wholesome as you can get. Mary Tyler Moore and Rhoda are basically nice girls and Archie Bunker would be just as upset about sex in commercials as respondents are. BUT there at the far right hand portion of the Table is a show that doesn't seem to belong, that makes no sense at all, that boggles the mind.

TABLE 4

Viewing Habits of Women Who Agree That "TV Commercials Place too Much Emphasis on Sex"



Over 40% of these women watch Soap Operas. Yet soap operas are filled with sexual material - bedhopping, infidelity, promiscuity, etc., etc. The attitude that there was too much sex on TV seemed totally incompatible with the behavior of watching soaps. It was at this point that we realized that we really had no idea as to what exactly people meant when they said that TV commercials have too much sex.

So we decided to stop looking at numbers and to start talking to some real live consumers. Six focus groups were conducted in September, 1977: three with blue-collar homemakers; three with white-collar homemakers. Women were also segregated by age: two groups were conducted with younger women, 18-25; two groups of middle-aged women, 30-40; and two groups with women 55+ years of age or older. All were married homemakers. Because our goal was to determine precisely what people were objecting to in terms of sex on TV, respondents were all prerecruited to agree with the statement, "TV commercials place too much emphasis on sex".

We began the groups by asking what TV programs they liked to watch. And sure enough, just as in the quantitative study, Family and soap operas were right up there with more wholesome TV shows. Why the anomaly?

Respondents were quick to point out there was no anomaly in their minds. They don't object to all sex on TV. They object to the wrong kinds of sex on TV. They like soap operas; they loathe Soap. Why the distinction?

They say sex is OK under certain conditions. It's all right when sex is implicit, not explicit. They rarely show anything in soap operas; you're saved the potential embarrassment by a quick cut-away. Sex is also all right when handled in a moral context. In soap operas, the heroes and villains are broadly drawn and easily recognizable. The sinners are eventually punished; the good guys are rewarded. They also say sex is always all right when you're prepared for it. And you're rarely more prepared than when watching a soap. Soap operas and Family also have the virtue of being shown at suitable hours, either the kids are away at school during the day, or safely tucked in bed at night.

Soap, on the other hand, is seen as being just the opposite. It doesn't show sex in a moral context; sex is treated lightly and deviant sexual behavior is condoned. Because it's immoral sex, respondents are far more likely to feel that any sexual nuance or innuendo is by default explicit.

How About Sex in Commercials

After exploring our respondents' attitudes toward sex on television we turned to the topic of commercials. We asked the women to give us examples of commercials that had "too much sex". The following commercials were mentioned most frequently: Noxema Shave Cream, Bic Lighter, Bic Razor, Ultra Bright, Muriel Cigars, Pearl Drops Underalls, Aviance, Shower Massage, Chicago Health Club, Tickle. Also mentioned frequently were a variety of commercials for "intimate" feminine products. After a general discussion about sex in commercials we showed the women four of these commercials (Bic Razor, Underalls, Muriel Cigars, Aviance Perfume) and explored their reactions in some detail.

By the end of our groups we felt we had a pretty good idea of what these women find offensive in commercials and why.

Sexual Innuendo

The Bic Razor commercial is a prime example of sexual innuendo. Here is a brief synopsis: It is a young man's wedding day. His sly uncle tells him to use Bic and get stroked in the morning (get it?) At the wedding his bride to be announces that he got stroked that very morning and now will get stroked every morning. The groom looks embarrassed, as well he might.

The young women in our groups characterized sexual innuendo as crude locker room humor. They found such suggestive insinuations embarrassing when others were present and offensive at all times. Any why shouldn't they feel this way - wouldn't everyone?

Well, not quite everyone. The older women saw nothing offensive in the Bic commercial and thought, in fact, it was kind of cute. When the moderator pointed out that the young women had found it offensive, the Edith Bunker syndrome took over. Older women would pause, look puzzled, frown their brows, and then say, "Ohhhhh... they couldn't mean..." Older women concluded only people with dirty minds would think that; younger women all had dirty minds. The only point about Bic that everyone agreed on was, "Well, at least it will go over the kids heads. They're not going to ask me about it".

Intimate Products and The Body

Consider this advertising for Underalls: Two girls are walking down the street away from the camera. The camera zooms in on their -uh- rear ends. One girl looks terrific. The other has panty lines and looks just awful. The girl with the problem tells us that underalls "make you look like I wish I looked."

In discussing what they hate in commercials women repeatedly mentioned that really intimate products should not be advertised on TV at all, and while panty hose may be considered less intimate than Kotex, it still constitutes an intimate product. In addition, women had already told us that certain parts of the body are inherently offensive and surely rear ends would fall into that category. We were therefore fully expecting our respondents to render scathing judgement upon this ad as it combined two inherently offensive elements - intimate products and bodily parts.

But in fact they liked it and didn't find it offensive at all. Sure the commercial focused on rear ends, but that was part of the product's reason for being. Panty lines are part of life; Underalls one way to get around them. As long as focusing on rear ends focuses the viewers attention on the product's primary benefit, it's not offensive.

The supposition that a focus on the product's primary benefit rather than on sex is important in mitigating offensiveness is supported by womens' reactions to another version of the Underalls ad. This particular ad ends with the line "Underalls makes me look like I'm not wearing nothing," and women hate it.

In addition to being grammatically incorrect, looking like one is not wearing nothing is not viewed as a primary benefit of buying Underalls. Women don't want to look like they are not wearing nothing and don't like people who do.

Male Fantasy Ads

What is a male fantasy ad? Well, let us give you an example. The ad for Muriel cigars is a male fantasy ad and goes something like this: A man sits working in his office alone at night and pulls out a cigar. A

super-sexy woman appears from nowhere and lights his cigar. She sings that he should let Muriel turn him on while she slithers around the office and ends up sitting on his lap.

You can always spot a male fantasy ad by the presence of two key elements: (1) foxy ladies, and (2) scenarios written by sexist piggy writers. The young women in our groups found this ad the most offensive of all the ads we showed them. Not only do they object to ads that place women in demeaning roles (lighting cigars, being at a man's beck and call) -- as any liberated woman would -- but in addition to that they find male fantasy ads personally threatening. Sexy women in ads make them wonder, "Is that how my husband wishes I looked?" This particular ad made them wonder, "Is that what my husband does at the office late?" Male fantasy ads contain unusually attractive women in unrealistic situations. It is this very lack of realism that young women find threatening and offensive.

But older women liked this commercial. Why? In part their consciousness hasn't been raised; it doesn't dawn on them that they should be offended. And even if the thought had crossed their minds they say, "Listen, honey. At my age anything that will get him excited just ain't half bad." In fact, one 64 year old grandmotherly type named Dorothy could hardly wait to get home and slink around her cigar-smoking husband singing, "Let Dorothy turn you on, that is my desire..." Moreover, they assert sex is always OK for traditional men's products. After all, they're advertising to men, not to them.

Female Fantasy Ads

If the Muriel ad is an example of a male fantasy ad then the Aviance ad is an example of a female fantasy ad: A beautiful young wife and mother is in the kitchen decked out in apron, rubber gloves - the whole housewife bit. Suddenly she begins a striptease while a female voice sings that she has had a full day of motherhood but is going to have an Aviance night. Scene shifts and she's in a bathtub full of Aviance bubble bath. Scene shifts and she's spraying Aviance perfume down her half opened blouse. Scene shifts to her fantastically handsome husband arriving at the door with flowers.

As it turns out young blue collar women hate female fantasy ads almost as much as they hate male fantasy ads. While they report that the bath scene and striptease are mildly offensive the main reason they hate this ad is because it portrays an unrealistic unattainable situation. They would all like to be in that situation but their lives just aren't like that. Their husbands don't come to the door with flowers like that and probably don't look like that either. Nor can they so conveniently get rid of their kids at night. They are unable to empathize with such female fantasy ads; instead they find them threatening and therefore offensive.

This time it was white collar young women who disagreed. The Aviance commercial was one of their favorites. They empathize with the commercial, and know just what it feels like to be in her situation. Maybe their husbands don't look quite like that, but they'd like them to. And they say it's good to show their children that "mothers and fathers love each other." The two older segments were split: some liked it, some couldn't relate, and some were just plain neutral.

Immoral Behavior

There is one thing that women of all ages agree on. Immorality offends them. The ads for both Bic Razors and Muriel Cigars violated our respondents' moral standards, which can be briefly summarized as follows: Physical

intimacy should only occur between consenting married adults (and even then only off-camera). In the Bic commercial he got "stroked" and they weren't even married yet. The Muriel ad was even worse - they weren't even engaged. At least the Aviance ad was about married people (and all of our groups made reference to this fact.)

Sexy Programs VS. Sexy Commercials

We think we've learned several things about commercials from this study. All offensive commercials are not created equal: they're not equally offensive, they're not offensive for the same reason, and they're not even offensive to the same people. But one still unanswered question occurred to us. Which is more offensive: sexy programs, or sexy commercials?

Even though some TV commercials are considered pretty offensive it made sense to us that TV programs would be considered even worse. For one thing TV programs last longer than commercials. If a commercial contains offensive material it will be over in a minute or less, but offensive programs go on and on and on. And besides sexual material is far more explicit on programs. And finally - how would any commercial possibly be as offensive as Soap which church groups have called: "leering sensationalism under the guise of comedy."

So we thought. Actually these groups were more offended by commercials. For one thing, there's no time for elaborate plot development -- in fact for almost any plot development in a 30 or 60-second commercial. Any sex used comes through that much clearer. And as a corollary they say you don't know sex is coming in a TV commercial. In TV programs, they build (and sometimes build and build) to sexual scenes; in commercials, wham bam, and it's over. Finally, they say sexy television commercials are harder to avoid. You know not to watch certain programs because you find them offensive. But commercials hit you out of the blue. By the time you get up to turn the channel, the offensive commercial is over.

Discussion

What do we think all this means? We think it means several things. First, sex is acceptable for certain products, products with no other reason for being than to increase allurements: cosmetics, perfume, and aftershave all fall in this category. Sex is also okay for advertising traditional masculine products, at least for non-liberated women: sex and beer, cars, liquor, and cigars mesh beautifully.

In addition, sex is permissible in female fantasies for young white collar women and male fantasies for older women. Sex is also all right when used in morally acceptable, eg. married, situations. And finally sex can be used when it's closely tied to the product's primary benefit.

We also learned that, for women at least, a commercial's offensiveness depends in part on the viewing situation -- on whether the viewer is alone or with others. Commercials are more offensive when others are present. Children giggle and ask difficult questions, husbands giggle and make offensive comments, and the presence of guests such as "my uncle, the priest" causes embarrassment. In the case of TV commercials, sex is more fun alone.

Finally, consumers, or at least these six groups, believe that television is an intrusive medium. Yes, you invite television into your home, but no, you don't want it to embarrass you. And even more than embarrass

you don't want it to foster immorality or to change existing sexual mores. Women aren't worried about changing their own mores. They're worried about people with impressionable minds, most specifically, their young children and all "those weirdos out there" who can be confused quite easily, and taught values the respondents don't believe in. As far as these women were concerned, TV not only mirrors social norms, it can change them. And that's potentially dangerous.

SEX ON TELEVISION, MORE OR LESS

William S. Rubens, National Broadcasting Company

Abstract

Are sex and violence substitutes for each other? There is no precedent or reason to think that they are. Yet they are frequently linked. While they are both emotional issues, they are distinct issues. The concern with television violence is derived from the possibility that it increases crime and delinquency. There is consensus in society that any such effects are undesirable. In contrast, there is no such societal agreement about the effects of television sex. Commercial television is the most conservative of the mass media when it comes to sex. There is no "real sex" on TV, only in-endo.

Some of the problems involved in discussing sex are well illustrated by a recent story in the New York Times. College biology students were asked to invite their parents to a seminar on pornography. After several pornographic films had been shown, one of the mothers volunteered that she had seen a porno flick before when her husband had brought one home. Shocked, her daughter the student blurted out "Mother!" Suddenly realizing what she had done, the daughter lamely explained to the class, "Well, after all, they are my parents."

When Dr. Wells asked me if I would be willing to serve on this panel about sex and TV, he also asked if I would address the question a lot of people have been asking lately: will sex fill the void created by the reduction of television violence?

As a researcher it was natural for me to think about the implications of the question before addressing myself to the answer.

This question is like the old saw, "Have you stopped beating your wife?" It seems to imply that sex and violence are functional equivalents in programming -- each one can simply replace the other. This assumption has little basis in fact. Also, the question as posed confuses two issues which must be kept separate: sex and violence. Just for the record, I want to state at the outset that I am opposed to the latter and in favor of the former. Finally, the question does not define what is meant by "sex" on television.

In my opinion, there is no sex on television. However, I've heard from people who watch the same programs I do who think they're saturated with sex.

I will cover some of these issues, try to answer the question that was intended, and discuss some research concerns I have with respect to sex on television -- and conclude with a discussion of network policy.

Sex And Violence: Functional Equivalents?

First, what about the assumption that sex and violence are substitutes for each other? There is no precedent suggesting that they are. If we look at theatrical movies, we can infer that sex and violence have escalated at the same time. Many movies show explicit sex and explicit violence as well. However this has not been the typical pattern in television. On TV, most of the action-adventure programs have little to do with sex and much of the so-called sex on TV occurs in comedies.

Furthermore, protests over television sex started some time ago when complaints about violence were also increasing. In other words, if those complaints reflect an increase in sex on TV, this increase took place before the reduction of action-adventure shows this season. Finally, the programming this season does not support the argument that sex takes over where violence leaves off. Action-adventure shows were replaced with all kinds of programs: new nonviolent dramas like "Man from Atlantis" and "Lou Grant," variety shows like "Redd Foxx," and all kinds of comedies -- including "Soap." Neither the new programs nor the networks' policies in replacing violent programs show any uniform pattern.

I think that all these developments show that there is no reason to think of sex and violence as functional equivalents. The content of television programs and the directions programming is taking are determined through complex processes which affect sex and violence independently.

Sex And Violence On TV: Two Different Issues

Sex and violence on television are emotional issues. Rational debate about them has been further impeded by the catch phrase "sex and violence" which lumps the issues together as if they were more or less the same. Indeed, the phrase is sometimes run together as though the two were one word: "sexanviolence." But for the broadcaster they are very distinct issues.

The concern with television violence is derived from the possibility that it increases crime and delinquency. The networks have reduced the amount of violence in their programming even though it has not been established that TV has actually affected crime or delinquency.

NBC maintains a panel of distinguished social scientists to advise Standards and Programming on potential harmful effects. Also, the networks' Broadcast Standards departments systematically supervise programming and attempt to eliminate portrayals of potentially harmful actions.

There is consensus in the society that any such effects are undesirable; no rational person would argue that there is merit in teaching or causing someone else to inflict harm on others and to act violently. As a result, broadcasters have the responsibility to act in a way which minimizes the chances of any such harmful effects.

In contrast, there is no such agreement about sex on television. Some people feel that sex is a legitimate topic for television to handle. Others feel that showing or discussing anything sexual on TV is in bad taste -- or "filthy." But the opinions of both groups are based on moral value judgments. And as far as perceived consequences are concerned, people generally agree that violence is anti-social and has harmful consequences, but there is no such consensus about sex.

In addition, there is no evidence that TV viewing affects sexual behavior and attitudes. There is little research on this issue, but NBC has collected some relevant data. As you may know, we are in the process of reporting on a 3-year longitudinal study on the effect of TV exposure on the behavior of children and teenagers.

The teenage portion included some questions on the boys' sexual behavior and attitudes. Therefore, we were able to analyze whether teenage boys who watched a lot of TV programs with sexual overtones of any kind were more likely to "neck," "go all the way," or have more liberal attitudes about such matters than boys who watched very little of these kinds of programs. We found that television viewing had no measurable effect on such behavior or attitudes. Even if we had found an effect there would be people who would argue for the philosophy, "make love, not war."

Since there is no moral consensus and since there is no scientific evidence of harmful effects, it is the responsibility of the broadcaster to analyze carefully what the controversy is all about. We can begin by asking what is this "sex on TV" that many people seem to get upset about?

What Is TV Sex Really Like?

Earlier this year, Eli Rubinstein and other researchers at the State University of New York at Stony Brook published the findings of the first content analysis addressing the issue of sex on television. After analyzing 61 prime time programs the authors reached this conclusion (I quote):

"The major finding was that physical intimacy appeared most often in less sensuous forms than one would expect from the public criticism of the portrayal of sexuality on current television programming."

If "sex on TV" means "showing people having sex," the finding was clear: there isn't any sex on TV. TV shows couples kissing and embracing, sex is discussed in a restrained manner, and there is innuendo -- but never actual sex.

These findings appeared to be surprising to the Stony Brook researchers; they aren't to me. There is little doubt that commercial television is the most conservative of the mass media when it comes to sex. Every variation of sexual activity has been shown in recent theatrical movies and has been described in recent books, and not far behind are Playboy, Penthouse, and Hustler, which rank among the nation's best-selling magazines.

If there is no "real sex" on TV, it is important to clarify what critics of TV "sex" are complaining about. One reason Rubinstein and his co-researchers expected real, sensuous sex on TV is that, as they point out, polls indicate that a clear majority of the American people think there is too much sex on TV and that many find "scenes of sex" more objectionable than violence on television. But the poll questions were as vague as the "Is sex going to replace violence" question: they do not clarify what people regard as "television sex," what they do and do not object to.

Opinions About Sex On TV

Critics of sex on television tend to assume that it is a firmly established fact that most Americans are critical of television sex. I deny that facts about attitudes concerning sex on TV have been established and I deny that the available data suggest that most Americans are critical of TV sex.

Among the few sources of information on this subject are a TV Guide poll, the mail we get, and public statements by organizations.

First, in April 1976 a TV Guide poll found that 58% of the respondents answered "yes" to the following: "In your opinion, is there too much emphasis on sex on TV?" Most researchers would agree that this is a classic example of the biased question. It uses emotional, value-laden words, its terms are vague and undefined, and it presents the respondent not with both sides of the issue, but with only one choice, tempting him into yea-saying. So the question is predestined to elicit a biased response and, in our judgment, is not an accurate gauge of public opinion on the issue.

However, even an accurately worded question would probably not be a good measure of opinions on this issue. There is hardly any subject more difficult to investigate than sexual behavior and attitudes, and it is unlikely that any one question can produce valid results.

This problem illustrates a favorite maxim of mine: Just because you ask a question and get an answer does not mean you have obtained information.

I can ask a question about virtually any topic. But it does not follow that the answers I get reflect genuine opinions on the issues I am trying to ask about. Surveys about sex attitudes probably fit this description better than almost any other topic.

The people who will not cooperate in a survey have different attitudes from those who will. This was a problem with the Kinsey report and has been a problem with all sex studies since, up to and including the Hite Report.

Consider, for example, Redbook Magazine's recent study on male sexuality, entitled "What Sex Means To The Man You Love." They printed the questionnaire in their magazine and asked their readers to hand the questionnaire to "the man you love" to fill out and send in. Mind you, the woman is not to peek at his answers.

Now let's look at some of the questions:

"How do you think your degree of sexual satisfaction compares with other men's? In the past few months who has been your main sexual partner? With how many other women have you had sexual relationships since your present relationship began? And finally: Please describe in some detail the most enjoyable sexual experience you've had."

Apart from the issue of self-selection, to believe that all 116 such questions get answered without underclaiming or overclaiming, strains my credulity. I guess this is what the statistician means by a "minimum likelihood" event.

However even when carefully done, the subject of sex presents special problems. For instance a study reported in Psychology Today found that a sample of college students -- usually considered a very liberated group -- experienced a surprising degree of blockage and embarrassment when questioned about the sexual activity of their own parents. As with the shocked girl in the New York Times story, the thought that their very own parents are still having sex seems to overwhelm some members of this presumably laid-back generation.

So as a researcher I have to remain highly skeptical over the ability of research to overcome the problems of response and non-response biases and extract the full truth on this emotional and guilt-laden subject of sex.

Viewer mail and public statements by organizations do not help much in an assessment of public opinion either. Organizations, whether small or large, are not representative of the public and neither is viewer mail. During the whole of last year, NBC received only 3,300 complaints about sex or morality on TV, less than 2% of the total letters received. Many of these letters are not about sex per se but are about sex in combination with sacrilege or with ridicule of prominent figures. This number also includes 900 letters from organized write-in campaigns.

It is interesting to learn what kinds of things these critics object to. There are very rare instances in which many letters and public complaints express dissatisfaction with a specific program or program segment. An example of this is criticism of a scene in an NBC movie, "Born Innocent," and the current furor over ABC's "Soap." In fact, numerous groups and individuals around the country were taking positions criticizing "Soap" even before it went on the air.

Most of the complaints, apart from these few specifics, are about everything on TV that could be construed as being related to sex: pretty girls in bikinis, jokes and innuendos in comedies and on talk shows, as well as dramatic treatment of topics such as rape. All these things are seen as "filth" and "permissiveness" on TV, which is perceived as endangering the "moral fiber" of this country.

But there are other voices. Some months ago I attended a seminar conducted by the Institute on Human Behavior dealing with sexuality and television. Basically it was the purpose of those running the conference to exhort the producers and networks' Program and Standards people to take a more forceful and perhaps more enlightened position in dealing with sexuality. The Institute urged more realistic portrayals of themes such as homosexuality, sex and the aged, and sex education for children.

It quickly became clear that the Hollywood creative community would be willing to incorporate these topics thematically into their programs. It became equally clear that this was not the problem. The resistance to frank portrayal of sexuality in TV rests with concerns of church groups, of affiliates, and of advertisers about offending American sensibilities and tastes, albeit minority tastes.

Thus the situation clearly reflects the lack of consensus about sexual issues in this society. The debate about sex on TV mirrors the society's own debates about the changing moral standards concerning sex. TV finds itself in the crossfire between opposing viewpoints.

TV Networks And Sex On TV

As sexual behavior and values have changed, television has changed. But it has not changed faster than society. TV is not and does not want to be a trendsetter of the sexual revolution. In fact, it lags behind the prevailing attitudes. And there is no reason to assume that the networks will alter that basic policy in the future. The reason is that this country's system of broadcasting has a built-in set of checks and balances which prevents programming from over-stepping the boundaries of what most people consider good taste. Programming decisions result from a complex interplay of many forces, which does not favor the emergence of radical trends.

Let's take a closer look at this interplay. The participants are the three networks with their program

departments, broadcast standards departments, and, of course, their managements. Then there are the producers, affiliates, and advertisers -- and last, but certainly not least, there is the viewing public.

The "viewing public" itself, as we have seen, is by no means of one mind on the issue of sex on TV. Attitudes toward sex, as exemplified in a Gallup question on "whether standards regarding the sale of sexually explicit material should be stricter than they are now" vary by age ... by sex... by education ... and by region of the country. For example, Gallup found 53% in the South favoring stricter standards, versus only 36% in the West. TV Guide's loaded question applying specifically to television, showed similar differences.

Reflecting these differences within the viewing public in the entire array of network affiliates -- themselves a vital element in the system of checks and balances governing television content. These affiliated stations are charged by the Communications Act with license responsibility for what they air. They represent a broad spectrum of political ideology, regional variation in taste and viewpoint, reflecting their close contact with their respective communities. As a general tendency, in the matters we're discussing here, we find there is commonly more concern about violence in the North-east, and more concern about sex in the South.

A second factor in the interplay of checks and balances is the force of competition among the three networks -- a factor which favors innovation. Today, one network appears to be giving viewers more cheesecake and making greater use of sexual overtones. Within the networks, we have program departments and broadcast standards departments with conflicting views over how innovative television should be. The people in these departments have different values and therefore often come up with different answers. The same is true of the network managements and producers.

A case in point is the different opinions about "Soap" mentioned earlier; it is a very complex question whether another network would have carried "Soap" or not.

From the first proposal of a program concept until the final edit of an episode, all these people -- writers, producers, directors, network management, programmers, and editors -- exert an influence. As a result, "Soap" done by another network -- and a whole different set of people -- would not have looked exactly the same.

But network managements not only take into account the views of their own executives and suppliers, they are also very sensitive to public reaction, including that of the critics and pressure groups. Similarly the advertisers, too, listen to the public, the critics and the pressure groups. In general, advertisers exert a conservative influence on the network management, since they do not want to see their product advertised in a program which might possibly annoy many viewers. Clearly, compromise is necessary to effectively balance the variety of views.

No wonder, then, that television has been slow to change with the times. The creative forces which press for innovation are nearly always restrained by conservative forces. Let me give an example: a few years ago, ABC was offered an innovative program. The network did not want to risk anything at that point so the program was offered to CBS and they scheduled it. "All In The Family" became a hit -- and also became the precursor of a host of programs with social messages, with the result that today we hardly think of "All In The Family" as the innovative program it was.

Other examples which justify network decisions to put innovative programs on the air despite opposition from some viewer groups include NBC's development of "Laugh-In," which was put on the air with the full fore-knowledge that some viewers might be concerned. There were indeed some protests, but the majority of the viewers welcomed the innovation, and incidentally "Laugh-In" became a hit. NBC's "Saturday Night Live" went through a similar process.

And we should not forget daytime serials. Our research indicates that most women viewers appreciate the treatment of all kinds of adult themes in daytime programming which they would not like to see during prime time, because in the afternoon they can relate to what they see on the screen in relative privacy.

Future Trends

It seems clear that in the current situation where there is no consensus about sexual attitudes and behavior, it is not possible for TV programmers to please everybody. The best policy under such circumstances is to try to steer a middle course. The networks have been doing just that, and they will not depart from that policy. However, TV will not go back to the standards of the 50's. It is the responsibility of broadcasters not to allow a minority of conservative critics to act as censors for the majority of viewers.

Indeed this is my main concern, the implied censorship issue raised by the ability of small pressure groups to influence what is seen by the American public. Even without direct formal censorship by the government, the pressure of these interest-groups can exert a chilling effect which in a subtler and more insidious manner produces the same consequences as formal, explicit censorship.

On the other hand, we have a responsibility to resist those who want television to become a leader in the sexual revolution. We have to continue being responsive to the majority of our viewers. If we don't, we invite government censorship. Especially in an emotional issue like this, the threat of government intervention is very real.

It is for all these reasons -- because of our responsibility to the viewer, the affiliate, the advertiser, because of our sensitivity to the minorities who criticize our programming, and because of our determination to avoid censorship -- that network television will change with the times, but on matters of sex will still remain the most conservative of media.

MEN AND WOMEN ON THE TUBE:
SEX AND SEX-STEREOTYPES ON TELEVISION

Susan Hesselbart, Florida State University

Abstract

Televised sexuality is examined in the context of gender stereotypes for three recent programs. Sex on television is usually implicit and--more important--unpleasant for television characters. Since the sexes are portrayed in stereotyped roles, even when the intention is "role-reversal", and often exploit each other, it is not surprising that televised sex is not much fun.

"Journalism is a lot more than blood and sex." Doris Day to Clark Gable in "Teacher's Pet". The outcry of concern over violence and sex on television today seems hardly to have progressed from "Teacher's Pet" nearly 20 years ago. However, this paper does not deal with the number of televised sexual episodes on television nor their explicitness. In fact, it is fair to say that there is very little explicit sex currently on television (e.g., Franzblau, *et al.*, 1977). Even on the program "Soap", roundly criticized for sexual explicitness before its premiere, there is little explicit sex--despite continual chatter from "Soap" characters on the topic. At this point, televised sexuality eerily resembles early 1960's movies and more explicit programs than "Soap" have already appeared on public television (e.g., "Poldark" during early 1977) with almost no averse publicity.

Rather, this paper concerns whether and how the "new adult" programs of the late 1970's portray sexuality as part of a gender stereotyped and segregated world. Obviously, sex is one of the most intimate relationships that can occur between women and men. However, sex on television can appear in different guises, from an affair deeply intertwined into the rest of the characters' lives to a "TV dinner quickie". Thus, the perspective here resembles content analyses of gender stereotypes in children's books (e.g., Weitzman, *et al.*, 1972), pornography (e.g., Smith, 1976), or everyday interaction (e.g., Henley, 1977). I am especially interested in how three programs during 1977 depart from or reinforce traditional gender patterns: the short-lived Norman Lear comedy, "All That Glitters"; ABC's "Soap"; and CBS' "We've Got Each Other."

These three programs were selected for several reasons. First, all are aimed at adults. The syndicated "All That Glitters" was usually shown in the late evenings, "Soap" carries messages about "parental discretion", while the themes of "We've Got Each Other" are aimed more at adults than children. Second, these programs were billed as a departure from traditional treatments of sex roles and/or sexuality. "Soap" has been designated the first prime time series to feature sexuality. "All That Glitters" was extensively advertised as a "sex-role reversal" comedy. "We've Got Each Other" has also been described as a role-reversal comedy (although less extreme than "All That Glitters"). Third, each series has featured some sexuality, from the explicit dialogue in "Soap", the constant sexual conquests in "All That Glitters" to the wedded couple on "We've Got Each Other". Thus in many ways, these programs have been viewed as harbingers of the future of sex and sex-roles on television.

Using these examples, this paper focuses on three main areas. First, the characters and settings will be described with reference to sex roles on television. The

second area concerns interpersonal relations between women and men. Finally, sex and gender are examined in reference to sex stereotypes in television characters and their interaction.

Gender Roles On Television

Certainly, the United States is a gender segregated society in many ways. Ample research documents polarized stereotypes of the sexes in all segments of the population, sex-typed labor markets, and segregated leisure activities (see Hesselbart, forth-coming, for a review). Nevertheless, television exaggerates this segregation--even when its intent is to do the reverse.

Several studies have addressed gender roles in the mass media. Unlike most American women in the late 1970's, few female television characters are employed, although, oddly enough, employed television women work at jobs largely occupied by men in "real life." On commercials, women are the voice--but not the voice--that sells products. Women are relatively likely to be victims, but not perpetrators of violent crime. And, women on television are younger and more physically attractive than their male counterparts (see, e.g., Isber and Cantor, 1975; Gerbner, *et al.*, 1977).

What happens when commercial television tries to be innovative in gender roles? As we shall see below, the "new" programs are still imprisoned by stereotypes.

"All That Glitters" explicitly stressed sex-role reversal. In this series, women were employed and men stayed at home. Within the work world, women were professionals and managers; employed men were secretarial and service staff. Men viewed for female attention, for marriage and for security, while the "masculinized" women ignored their families and thought only of business and sexual conquest. After only a few months, this series folded. According to some television critics, the viewing public was not ready for a role-reversal comedy.

However, "All That Glitters" only changed the surface of gender-stereotyping. Stereotypically, husbands or boyfriends were usually older than their wives or girlfriends. Women executives acted short-tempered and bitchy. Women mused in bars for long scenes about the characteristics and relationships of their boyfriends--again more in line with feminine than masculine stereotypes. And, as discussed below, the more subtle aspects of interpersonal dominance between the sexes remained unchanged.

"We've Got Each Other" is a "softer" version of role-reversal. Judy Hibbard is an office manager for a temperamental photographer. Her husband, Stewart, is a "Househusband" who cooks and cleans while Judy goes out to work. However, Stewart is employed (writing mail-order catalogues), and, moreover "takes Judy out to eat". Meanwhile, Judy is the emotional one, losing her temper easily and concerned about her attractiveness to men (Stewart, meanwhile, feels he is already attractive to women).

"Soap" makes no pretense at role-reversal, but, if anything, emerges as more insulting to both sexes than most of prime time television. First, nearly every

character is incompetent.¹ Seemingly wealthy Chester Tate is poor at business and stoops to shady practices. His wife is easily deceived about Chester's affairs, her daughter's affairs, and her son's pornography collection. Women are generally stereotyped as peabrains, manipulators, or nymphomaniacs (one exception is Mary Campbell). The men are reduced to throwing food (which the women clean up). Homosexuals are stereotyped as trans-sexuals desiring sex-change operations and Mary's son is a drag queen who looks better than his mother in dressess.

My point is not that television should have programs on role-reversal or sexuality. Rather, it is that these "departures" from traditionalism are trapped by stereotypic writing and directing. For example, men on "All That Glitters" usually portrayed the most negative female stereotypes--frumpy housewives, dilettantes, sexual manipulators. Female characters did not portray any more flattering portraits of men--Don Juans, unconcerned fathers, or heartless businessmen. Reverse such stereotypes back to their traditional gender and the characters on "Soap" emerge. Even on "We've Got Each Other," more gentle in tone, Judy is a totally incompetent cook who burns toast and opens bacon with her teeth (a common stereotype of men at home) and Stewart's work is not to be taken seriously.

Television does not yet seem capable of portraying an "androgynous" world (e.g., see Bem, 1975) in which characters combine both desirable male and female traits. Versions of "role-reversal" which tacitly accept traditional gender stereotypes appear to be television's answer to greater sex-role egalitarianism. We shall see below that dominance patterns between the sexes are also traditional, again, even when "role-reversal" is the program topic.

Dominance And Gender On Television

In American society, the sexes are not only considered to be very different from each other, but men are usually expected to dominate women. Domination can be explicit, as in the male commercial characters who tell women how to do the wash or when women are employees but rarely the boss. However, more subtle dominance relations are also expressed in mixed-sex interaction.

Researchers such as Henley (1977) focus on nonverbal gestures that convey dominance. For example, status superiors (employers, parents) touch status inferiors (employees, children) far more frequently than the reverse. Setting conversational topics or interrupting others also signifies status. Superiors are more likely to "first-name" inferiors than the other way around.

Dominance patterns on programs such as "All That Glitters" provide further insight on why "role-reversal" television never quite seems realistic. While who earns the money or who keeps house may be gender-interchanged, more subtle patterns of sex-status continue. The result is jarring to those who consciously or nonconsciously are aware of dominance between men and women.

Take, for example, "All That Glitters." Continually, male secretaries casually touched their female bosses, while women touched men only in love, sex, or familial situations. Men interrupted women or introduced new conversational topics. While women sometimes lost interest, they tolerated these intrusions of personal space in a way that seldom occurs in traditional gender interaction (e.g., see West, 1977).

In "We've Got Each Other," it is Stewart who criticizes Judy's appearance or who persuades Judy to return to the job she quit. As in "The Mary Tyler Moore Show," Judy's boss calls her by her first name, although he is

"Mr. Jerome." However, it is interesting to note that Judy is slightly taller than her husband--a nonverbal detail that lends greater credence to this program's claims of "role-reversal."

Again, the argument is not that television should have more series on "role-reversal" (or even androgyny). However, the experiments tried to date are internally contradictory. If women are the status superiors ("All That Glitters", "We've Got Each Other"), they should be portrayed that way, both in demographics and in interpersonal interaction. Otherwise, the audience is at least subliminally aware that the characters and roles do not "fit."

Despite attempted innovation, television has fared poorly in its attempts to deal with traditional sex-roles. Stereotyped characters and dominance patterns continue. It should come as not surprise that such trends also occur when television deals with sexuality. Since sexuality is one arena where cross-sex interaction is nearly a cultural mandate, these stereotypes become important in the rather melancholy portrayal of sex on television.

Sex And Sex Roles On Television

An important issue in television violence concerns what viewers learn from filmed hostility. For example, increased television viewing correlates with a "jungle world" perspective in which others are untrustworthy (e.g., Gerbner and Gross, 1974; Gerbner, et al., 1977). The debate on television aggression and subsequent viewer aggression is too long and well known to reproduce here. Give such attention to televised violence, it seems reasonable to wonder what lessons viewers receive about sexuality.

First, viewers certainly do not receive a sex manual. As mentioned earlier, televised sex is far more implicit than explicit. The programs discussed here are not exceptions. At most, characters embrace or are shown adjacent in bed. While "Soap" characters discuss sex, they do so in vague terms (e.g., "that glowing feeling" or "his you-know"). "Soap" hardly constitutes an action comedy. Sex is generally unpleasant on these "new adult" comedies. Even while "Soap's" Jessica glows from her affair with a tennis instructor (who is also her daughter's lover), she is filled with guilt. Her attempts to confess to her husband fail, since he is engrossed with simultaneously juggling several affairs and is being blackmailed by his secretary (and ex-lover). Jessica's sister Mary must cope with an impotent husband who blames her for the impotency. The sexual encounters on "All That Glitters" were similarly joyless. Male "sex objects" worried about their attractiveness, propriety, or "keeping score", while the female "aggressors" faced either one night stands or troublesome lovers. Sex is far from fun on television comedies.

Third, sex on these programs is usually a tool. Sex is used for blackmail; material gain, manipulation, or to reassure the participant about his/her physical attractiveness. Thus, sex is done to (or from) but rarely with others. Given such motives, it is no wonder that no one looks very happy.

On Television comedies, sex also seems to provide one of the few opportunities for cross-sex interaction. In "All That Glitters", (besides affairs) cross-sex interaction was restricted to employer-employee relations or arguments between spouses. "Soap" is much the same. Outside of sex, blackmail, or spats, the sexes do not work, chat or play together. "We've Got Each Other" provides a partial contrast: Judy and Stewart are friends: they are able to convivially discuss everyday

affairs as well as lounge in bed in identical pajamas.

To summarize then, televised sex is usually implicit and apparently not much fun for the participants. It is hard to see how the situation could be otherwise when sex is a manipulative strategy (often used by women or "role-reversed" men to gain power) and the sexes have little common ground (which would appear inevitable given the stereotyped characters). These are lessons to viewers that many of us deplore. But, certainly, televised sex appears to be far from the hedonistic orgies envisioned by some television critics and pressure groups.

Conclusions

"All That Glitters," "Soap," and "We've Got Each Other" were selected for discussion in this paper because each, in its own way, has been conceived as an innovation in television programming vis a vis sexuality and gender roles. Despite attempts at innovations, however, these programs are old wine in new bottles: traditional sex stereotypes and patterns of gender domination continue. The sex-typed inelasticity of the characters and their relations spill over into televised portrayal of sexuality. In a world in which women (or role-reversed men) must worry about security and lack efficacy, using sex for manipulation seems logical. When women and men are portrayed as so different as to make friendship unlikely, joyous sexuality is also unlikely.

What can we expect in the future of television? Out in the "real world," the labor force participation of women and men is converging, people are divorcing more, and marrying and reproducing less. In its fashion, television is attempting to reflect such changes.

Whether or not "role-reversal" comedy or more explicit sexuality become television staples depend on how well these concepts "sell." "All That Glitters" had a short season and "We've Got Each Other" is doing very poorly in weekly ratings. In this paper, I have suggested that unattractive characters may be as important in the low success of these programs as the "role reversal" concept. In addition, the inconsistencies between concept and execution in these two programs renders the characters and situations unbelievable. "Soap", on the other hand, is currently a success, although a recent article stated that an attempt will be made to render the characters more three-dimensional.

It is a shame that television has done so comparatively little with the concepts of sexuality and sex roles. Both are certainly important issues and could be well-treated in humor. How a couple treats sexuality in the context of a caring relationship, or how to raise a family in a time of changes are two examples of topics that television comedy could address. But, with notable exceptions, prime time broadcasters have expended little imagination to explore such issues. In the future, I hope that they will try.

References

- S.L. Bem, "Sex Role Adaptability: One Consequence of Psychological Androgyny." Journal of Personality and Social Psychology, 1975, 31: 634-643.
- S. Franzblau, J.N. Sprafkin, and E.A. Rubinstein, "Sex on TV: A Content Analysis." Journal of Communication, 1977, 27 (2): 164-170.
- G. Gerbner and L. Gross, "Violence Profile No. 6: Trends in Network Television Drama and Viewer Conceptions of

Social Reality." Working paper for the Anaheim School of Communication, Pennsylvania State University, December, 1974.

G. Gerbner, L. Gross, M.F. Eleey, M.Jackson-Beeck, S. Jeffries-Fox, and N. Signorielli, "TV Violence Profile No. 8: The Highlights." Journal of Communication, 1977, 27 (2): 171-180.

N. Henley, Body Politics. Englewood Cliffs, N.J.: Prentice-Hall, 1977.

S. Hesselbart, Women and Men: Today, Yesterday and Tomorrow. Boston: Little, Brown, in progress.

C. Isber and M. Cantor, Report of the Task Force on Women in Public Broadcasting. Corporation for Public Broadcasting, 1975.

D.D. Smith, "The Social Content of Pornography." Journal of Communications, 1976, 26 (1): 16-24.

L.J. Weitzman, D. Eifler, E. Hokada, and C. Ross, "Sex-Role Socialization in Picture Books for Preschool Children." American Journal of Sociology, 1972: 1125-1150.

C. West, Communicating Gender: A Study of Dominance and Control in Conversation. Ph. D. Dissertation, University of California, Santa Barbara, 1977.

Footnote

- ¹Incompetence may, however, simply be a characteristic of most characters in television situation comedies.

ENERGY CONSERVATION AND TRAVEL BEHAVIOR

Jeffrey S. Milstein, Ph.D.
Department of Energy
Washington, D.C. 20461

Abstract

Americans use one quarter of all energy consumed in the United States in automobile passenger travel. This paper describes the ways Americans could reduce this energy consumption, and, using empirical data, describes the extent to which these energy-conserving behaviors are or are not being practiced.¹

The Energy Problem

Today, with only 5 percent of the world's population, Americans are consuming 30 percent of the world's petroleum production.² This consumption of oil comes at an increasingly high cost, for this Nation must import an increasingly larger share of its oil. Imports for the first half of this year amount to 47 percent of total consumption, compared to 39 percent for the first half of 1976. The cost of this imported oil is growing at a staggering rate: this year alone U.S. citizens will pay foreign oil producers \$45 billion for the oil we are importing! This is up from \$37 billion in 1976, \$27 billion in 1975, \$8 billion in 1973, and only \$4 billion in 1971. Moreover, 41 percent of these imports come from countries that have shut off the flow of oil to us in the past, and could again.³ The potentially enormous negative consequences for our economy if oil is again shut off, as well as the consequences for our national security and independence of our foreign policies, are urgent national and international reasons for the citizens of our Nation to consume less energy.

Travel by Car: A Great User of Energy

What has this to do with travel behavior? Passenger travel is an enormous user of our energy, particularly of petroleum. Sixteen percent of all our energy is used directly in passenger travel. Nine-tenths of all passenger travel is in automobiles. Travel by car uses 57 percent of all energy used in transportation in the United States. (Air passenger travel is a distant second, using 6 percent of U.S. transportation energy.)

About 2 out of every 5 BTU's of energy used in automobile passenger travel is used indirectly. Thus, to the 14 percent of all our energy directly consumed by automobiles as gasoline and oil, one must add such energy use for automobile travel as manufacture of motor vehicles, parts, tires, and repairs (about 4 percent more of our total energy use); highway construction (about 2 percent); and petroleum refining and distribution losses (about 5 percent). Thus, directly and indirectly, automobile passenger travel uses one-fourth of all our Nation's energy and

an even greater percentage of our oil--about 30 percent.⁴ American cars are now using about 1 out of every 10 barrels of oil produced in the world! As a major consumer of energy then, automobile travel by Americans is a major focus for energy conservation measures.

The Department of Energy and its predecessor organizations have recommended for almost 4 years various ways that private citizens and consumers can reduce their energy use. In the area of transportation behavior, I will go through those major recommendations and then describe (using empirical data) the extent to which the American public is or is not practicing these ways to conserve energy.

Buying Efficient Cars

The first travel-related behavior that saves energy is for people to buy efficient cars that suit their needs. In 1977, the average miles per gallon of new cars sold, including imports, was 18.6 miles per gallon. That was up from 17.6 in 1976, 15.6 in 1975, and 13.9 in 1974.⁵ Thus, from 1974 to 1977, there has been a 34 percent increase in new car fuel efficiency. This has come about as car manufacturers have been required by law to increase the average efficiency of new cars.

Related to the operating fuel efficiency of new cars is their size. Smaller cars also use less indirect energy, such as that necessary to produce additional amounts of steel for larger cars. The market shares for different sizes of cars, including imports, in 1977 and 1976 are as follows:⁶

	1977	1976
Subcompacts	25%	23%
Compacts	22%	25%
Intermediates	28%	28%
Full-size	24%	23%
Vans	1%	1%

Note that part of the increase in subcompact sales from 1976 to 1977 includes an increase in foreign car sales from 14 percent of the total market in 1976 to 19 percent in 1977. Nine out of ten foreign cars are subcompact

⁴"National Transportation: Trends and Choices," U.S. Department of Transportation, January, 1977, p. 33.

⁵Data are from the Motor Vehicle Manufacturer's Association.

⁶These data are from January 1 to August 20, for the respective years. The data comes from Wards Communications, Inc.

¹The findings and interpretations in this paper should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Department of Energy or of the U.S. Government.

²Federal Energy News, FEA, Aug. 29, 1977.

³Federal Energy News, FEA, Aug. 29, 1977.

in size. Thus, part of the gain in the new car fleet's fuel efficiency is due to the increased sales of foreign cars. While importing cars hurts our balance of payments, it does not also make the United States strategically vulnerable as does the importing of oil.

Ridesharing

Ridesharing, especially to and from work, but also including shopping with a neighbor, is one of the best ways that American consumers could cut down on their vehicle miles traveled and therefore save energy. A three-person carpool uses just one-third of the energy that the three people driving alone would use. About one-third of all private automobile mileage is for commuters to and from work. If the average occupancy (currently 1.3 people per commuter car) were increased by just one person, not only would each commuter reduce his cost, energy use, and driving stress, but the Nation would save more than 700,000 barrels of gasoline per day.⁷ This is approximately one barrel out of every ten for gasoline that Americans use. If just one gallon of gasoline were saved each week for each automobile in the country, we would save about 5.2 billion gallons of gasoline in a year or about 7 percent of all the demand created by all of our passenger cars.

The empirical data shows that half of all trips have only one occupant per car. Of trips to and from work, approximately three out of four were in one-occupant cars.⁸ A 1973-74 national survey of work trips shows that approximately one out of four persons is going to and from work in a vehicle with two or more people.^{9, 10}

Though these data come from studies by the Department of Transportation in 1972 and 1973, more recent FEA surveys I have directed show that carpooling has not significantly increased since the oil embargo of 1973-74.

Carpooling varies tremendously around the country: as many as half the people in Washington, D.C., and in Minneapolis-St. Paul carpool to work. Moreover, where special carpool facilities exist, such as the Shirley Highway bus and carpool express lanes into Washington, D.C., from Northern Virginia, we find as many as three out of five people in these corridors taking cars are in carpools. But on a national basis, the great energy savings potential of carpooling and vanpooling is far from being realized.

Public Transit

Public transit uses just one-sixth to one-third the energy per passenger mile that single occupancy cars do.¹¹ How-

⁷"Tips for Energy Savers," Federal Energy Administration, 1977.

⁸"Nationwide Personal Transportation Study, Automobile Occupancy" Report No. 1, April 1972, DOT/Federal Highway Administration.

⁹"Nationwide Personal Transportation Study: Home to Work Trips and Travel," Report No. 8, August 1973, DOT/Federal Highway Administration.

¹⁰Note that the distribution of vehicles is somewhat different from the distribution of persons because of multiple occupancy in carpools.

¹¹"Buses Hailed, Rail Transit Scored in Energy-Use Study by Congress," New York Times, Sept. 23, 1977, p. 18.

ever, various studies show that only 6 to 8 percent of the public takes public transportation to and from work.¹²

In the FEA surveys that I have directed, an analysis of why people do or do not use public transit for going to and from work shows that 17 percent of the total public say that mass transit does not run at the right times, 11 percent say it does not go where they need to go, 9 percent say it does not run often enough, 8 percent say it takes longer, 5 percent say that they do not because public transit encounters delays and holdups, 4 percent cite overcrowding on public transit, and 1 percent cite dirty vehicles as the drawback from their point of view.¹³

In a multivariate analysis of the FEA survey data, we found that half of the users of public transit encountered heavy traffic and jammed highways on the way to and from work, whereas only 27 percent of the drivers of their own car did so, and 38 percent of carpoolers did so. We infer from this that traffic jams or heavy traffic are important inducements for people to take public transit. An analysis of the 30 percent of our total population who have a choice between taking cars or public transit to and from work (i.e. have public transit available to them) shows that the factors people seem to consider in taking public transit are: the distance and duration of travel, the convenience and timing of the schedules, the proximity of the transit line to where they work and live, the cost of the travel, and whether or not they encounter heavy traffic or traffic jams on the way to and from work.¹⁴

Analysis of the FEA survey data also show that a large fraction of the people who take public transit are also people who would carpool and vice versa. These people are different in their attitudes and behavior from people who drive alone to and from work. An important implication from this finding is that promoting public transit ridership might be at the expense of some carpooling, and the promotion of carpooling might be at some expense to public transit ridership. Again, the enormous potential for energy savings in public transit riding is not being realized.

Walking

Five percent of the public walks to and from work. The percentage who use this energy saving mode of transportation appears small, but is nearly as large as the percentage who take public transportation to and from work.¹⁵

Vacation Travel

In a recent survey, it was found that six out of ten households in the United States took one or more vacation

¹²"Nationwide Personal Transportation Study: Home to Work Trips and Travel," Report No. 8, August 1973, U.S. DOT/Federal Highway Administration; also, FEA survey by Opinion Research Corporation, 1975, Vols. 11 & 12.

¹³"The Public's Attitudes Toward and Knowledge of Energy Related Issues," Highlight Report Vol XI, FEA, June, 1975.

¹⁴"General Public Behavior and Attitudes," Highlight Report Vol. XII, FEA, July, 1975.

¹⁵"Consumer Attitudes, Knowledge and Behavior Regarding Energy Conservation," FEA, December, 1976.

trips in 1976, and half of all households took a vacation trip by car. The average number of trips taken by vacationing households was more than two trips per year. Half of the households who take vacation trips take two or more vacation trips per year. On vacation trips, the average travel party consists of three people (typically a family). The average mileage driven per year is about 3,500 for all household vacations. (This is about a quarter of all miles driven per year.) The median income of families taking automobile vacation trips in 1976 is about a quarter higher than the median income for all families in the United States (i.e., \$17,640 versus \$14,000 per year).¹⁶

Vacation travel took a temporary dip during the Spring of 1974 during the oil embargo, and during the 1975 recession. Vacation travel increased, however, in 1976 and again in 1977. Thus, energy-conserving travel behavior such as vacationing near home, staying in one place instead of traveling around, and taking a train or bus instead of the family car, have not been significantly practiced since the oil embargo and subsequent recession. Americans continue to use a large amount of energy on their vacation travel--an integral part of the lifestyles of a majority of Americans.

Driving Slower: The 55 Mile Per Hour Speed Limit

Most automobiles get about 20 percent more miles per gallon on the highway at 55 miles per hour than they do at 70 miles per hour. Fatal automobile accidents are also less likely at the slower speeds. Thus there are both energy and safety reasons to drive slower on the highway.

In 1973, the mean speed on the highways which had free flow conditions (i.e., the 45 percent of American roads which have only the 55 mile per hour speed limit as a control) was 60 miles per hour with a standard deviation of 10. In 1974, the mean speed dropped to 55 miles per hour with a standard deviation of 6.17

In 1977, the majority of States have average speeds between 55 and 60 miles per hour. This shift in the mean speed that occurred between 1973 and 1974 (when the 55 mile per hour speed limit became law in every State) has continued since 1974. The proportionate gasoline savings have also continued as a result of people driving at this more efficient speed on the open highway.¹⁸

Other Efficient Driving Behaviors

There are other driving techniques that a careful driver can use to get 20 percent more miles per gallon than the average driver and 50 percent more than a wasteful one. These techniques include accelerating smoothly and moderately, driving at a steady pace, avoiding stop and go traffic, minimizing braking, not letting the motor idle for more than a minute, not overfilling the gasoline tank,

¹⁶"The 1977 Automobile Travel Intention Survey" December 2, 1976, unpublished paper by James J. Gibson, 3M National Advertising Company.

¹⁷"Transportation Safety Information Report," Assistant Secretary for Environment, Safety and Consumer Affairs, U.S. DOT, 1977.

¹⁸U.S. DOT/Federal Highway Administration, Highway Statistics Division, "Quarterly Speed Summary," June 1977.

planning trips carefully, having the car tuned (3-9 percent savings), keeping the engine air filter clean, checking tire pressure regularly (2 percent savings per pound of underinflation), using radial tires (3-10 percent savings) and removing unnecessary weight from the car.¹⁹

Reliable and valid data on these kinds of behavior on the part of drivers is not readily available. But it is a part of travel behavior where energy savings can add up to significant amounts.

The Bottom Line: Gasoline Consumption

Distilling all automobile travel behavior down to the bottom line of gasoline consumption, we find that motor gasoline demand for January through August of 1977 averaged 7.1 million barrels per day, or 81 percent of our imports. That is nearly 300 million gallons of gasoline per day. While that averages only 2 2/3's gallons per automobile per day (or an average of about 40 miles driven per day), it is an amount that is 2.1 percent higher than the average for 1976.

Since 1972, gasoline consumption has been increasing at an average annual rate of 2.6 percent per year. However, new car registrations have been increasing an average of 3.3 percent per year from 1972 to 1976,²¹ which is slightly faster than gasoline consumption. Thus, even though in the aggregate, Americans are using more gasoline in cars, on a per capita basis, there has been a slight amount of energy conservation in gasoline used by the American people. (One explanation for this is that in the growing number of two-car families, the second car is not driven as much on long family vacation trips.)

Conclusion

In summary, Americans appear to be making some progress in conserving energy in such areas as buying more efficient cars and driving slower on the highways, but little or no progress in such other important areas as carpooling, using public transit, and reducing vacation travel. The challenge that lies before us still is to reduce by significant, rather than trivial amounts, the fuel consumption of Americans as they travel. We must use all of the means possible to save energy. We cannot forget or ignore the enormous international and domestic economic and political stakes that are present in this challenge.

References

Federal Energy Administration. "Consumer Attitudes, Knowledge and Behavior Regarding Energy Conservation." Prepared for FEA by Opinion Research Corporation, Princeton, New Jersey. (National Technical Information Service: Springfield, Virginia, December 1976). (FEA/D-76/469)

Federal Energy Administration. Federal Energy News. (Washington, D.C., August 29, 1977).

¹⁹"Tips for Energy Savers," FEA, 1977.

²⁰"Monthly Energy Review," FEA, August, 1977.

²¹"Total Automobile Registrations," U.S. DOT, 1977.

Federal Energy Administration. "General Public Behavior and Attitudes Regarding Vacation and Business Travel, Beverage Containers, Reasons for Using Mass Transit." Highlight Report, vol. 12. Prepared for FEA by Opinion Research Corporation, Princeton, New Jersey. (National Technical Information Service: Springfield, Virginia, July 1975). (PB 244 969/AS)

Federal Energy Administration. "Monthly Energy Review." (Washington, D.C., August 1977).

Federal Energy Administration. "The Public's Attitudes Toward and Knowledge of Energy-Related Issues." Highlight Report, vol. 11. Prepared for FEA by Opinion Research Corporation, Princeton, New Jersey. (National Technical Information Service: Springfield, Virginia, June 1975). (PB 244 987/AS)

Federal Energy Administration. "Tips for Energy Savers." (Washington, D.C., 1977).

James J. Gibson. "The 1977 Automobile Travel Intention Survey." 3M National Advertising Company. (unpublished report, December 2, 1976).

Holsendolph, Ernest. "Buses Hailed, Rail Transit Scored in Energy-Use Study by Congress." (The New York Times, September 23, 1977). p. 18.

Department of Transportation. "National Transportation: Trends and Choices." (U.S. DOT, Washington, D.C., January 1977). p.33.

Department of Transportation. "Nationwide Personal Transportation Study Automobile Occupancy." Report No. 1. (DOT/Federal Highway Administration, Washington, D.C., April 1972).

Department of Transportation. "Nationwide Personal Transportation Study: Home to Work Trips and Travel." Report No. 8. (U.S. DOT, Washington, D.C., August 1973).

Department of Transportation. Assistant Secretary for Environment, Safety and Consumer Affairs. "Transportation Safety Information Report." (U.S. DOT, Washington, D.C., 1977).

Department of Transportation. "Total Automobile Registrations." (U.S. DOT, Washington, D.C., 1977).

Department of Transportation, Federal Highway Administration, Highway Statistics Division, "Quarterly Speed Summary." (U.S. DOT, June 1977).

CONSUMER ORIENTED TRANSPORTATION PLANNING:
AN INTEGRATED METHODOLOGY FOR MODELING
CONSUMER PERCEPTIONS, PREFERENCE AND BEHAVIOR

Alice M. Tybout, Northwestern University
John R. Hauser, Northwestern University
Frank S. Koppelman, Northwestern University

Abstract

This paper presents research examining a model of consumer transportation behavior which integrates system characteristics, perceptions, preference, choice and situational constraints. The research findings are used to develop and evaluate strategies designed to increase consumer acceptance of public transportation.

Introduction

Public officials and private citizens alike acknowledge the need for local public transportation. Public transportation is viewed as necessary both to provide mobility to individuals who do not have any private means of transportation and to offer an alternative to private transportation for all individuals. In addition, public transportation is perceived to be desirable because it offers advantages over private transportation in terms of energy efficiency, environmental pollution, and congestion.

However, several observations indicate that individual and community needs are often not well met by existing public transportation systems. Most local transportation services require substantial operating and capital subsidies, while at the same time, have considerable excess capacity, particularly during off-peak hours. Furthermore, the low utilization of public transportation and the attendant prevalence of private auto trips has resulted in traffic congestion, parking problems, and environmental pollution in many communities, and has contributed to national gasoline shortages.

If strategies are to be developed to gain consumer acceptance of public transportation, a clearer understanding of consumer transportation behavior is required. Specifically, research must be undertaken to:

- (1) identify and measure consumer perceptions of transportation alternatives and examine how such perceptions are developed. (i.e. their relationship to actual system characteristics),
- (2) determine the relationship between consumers' perceptions and their mode preference and choice,
- (3) provide a strong test of the relationships established in (1) and (2). This can be done by developing and implementing strategies based on the knowledge accumulated in (1), (2), predicting the impact of these strategies on perceptions, preference and choice and then evaluating their actual impact on these factors.

In this paper the preliminary results of research undertaken to address the above issues is reported. First the relevant literature is reviewed. Then, the research method and results are summarized. Finally, the practical and theoretical findings are discussed and plans for future research are outlined.

The Literature

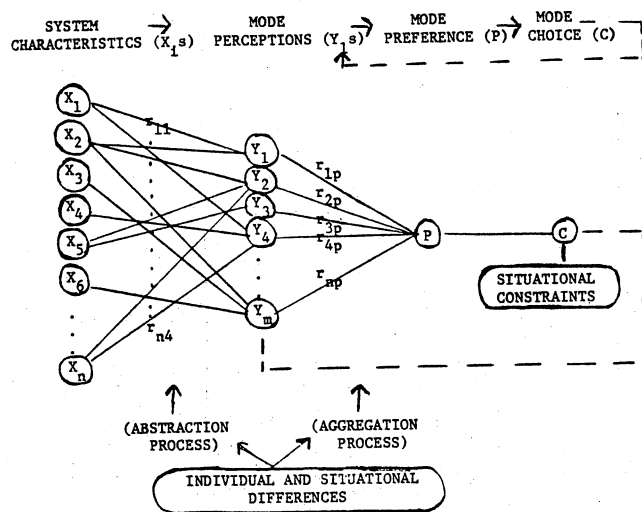
Efforts to understand consumer transportation behavior began in the late 1950's with aggregate studies correlating system characteristics (e.g. travel time, frequency, cost etc.) and community characteristics (e.g. income, education, density etc.) with demand for transportation alternatives (Martin, Memmett, and Bone, 1961; McLynn and Woronha, 1969; Quandt and Banmol, 1966, SARC 1963). Although these models performed well in special circumstances, they did not adequately represent consumer behavior and they gave little guidance for the development of strategies to directly influence such behavior. As a result, disaggregate demand models were developed in the early 1970's to examine the relationships between these variables and mode choice on the level of the individual consumer (Ben-Akvia, 1973; Charles River Associates, 1972; Koppelman, 1975; McFadden, 1970). These disaggregate models, which are still widely used today, lead to a clearer specification of the relationship between system characteristics, consumer demographics and mode choice than did aggregate models. However, like the aggregate models, they concentrate on system and community characteristics. Thus, they fail to provide an understanding of consumers' transportation decision making processes and are not sensitive to the wide set of strategies that can be developed to influence consumer behavior without expensive changes in system characteristics. Additional research is required to gain a better understanding of the consumer and make available the full range of strategic opportunities to community transportation planners and managers.

The critical factor present in consumers' decision making but absent in traditional demand models is consumers' perceptions, which mediate the relationship between system characteristics and mode choice. In recent years several transportation researchers have acknowledged the importance of perceptual variables (e.g. convenience, comfort) and have focused on quantifying these variables so that they can be included, along with system characteristics such as time and cost, in disaggregate mode choice models (Nicolaidis, 1975; Spear, 1976). While this effort to quantify perceptions is important it is not sufficient. An adequate understanding of the relationship between perceptions, system characteristics, preference and choice is also needed.

One way of viewing the interrelationship between system characteristics, perceptions, preference and choice is given by the model of consumer transportation decision making in Figure 1. This model, which is an extension of Brunswick's lens model (see Brunswick, 1952, Hammond, 1966), is similar to models which have been used to describe consumer response to new products/services (Green and Wind, 1975; Hauser and Urban, 1977; Pessemer, 1977). In the model, the system characteristics (x_i 's) serve as cues used by the consumer in forming his perceptions (i.e. evaluation of convenience, safety, comfort, liking etc.) of the various modes. Each system characteristic is an imperfect indicator of any particular perception and the degree of the association between

these factors is represented by the r_{ij} s. Thus, travel time may serve as a partial indicator of convenience for a particular mode. Furthermore, a system characteristic such as travel time may influence several different perceptions in different ways (i.e. travel time may be negatively correlated with convenience but positively correlated with safety). This process of using system characteristics as cues in forming perceptions is called abstraction. Once perceptions are formed they are aggregated to determine choice. A degree of association between the perceptions and preference is represented by the r_{ip} s. Preference, tempered by situational constraints such as mode availability, in turn, directs choice. Finally, choice and experience may feed back modifying mode perceptions. All situational and individual differences not represented in this basic model influence decision making by influencing the manner in which the individual forms perceptions or aggregates perceptions to direct choice.

FIGURE 1
A MODEL OF CONSUMER TRANSPORTATION BEHAVIOR



There are several practical implications of this model. First, if the model accurately represents consumers' transportation decision making, then researchers should concentrate their efforts on understanding the abstraction and aggregation processes (i.e. are linear or non-linear models used etc.) and should not confound these two processes by using both system characteristics and perceptions in one model predicting preference or choice since these variables are likely to be highly correlated. Nor should they develop mode choice models from system characteristics directly. Second, because the model represents the stages in the consumer decision making process it can help the planner or manager diagnose problems in the transportation system. Problems may either be the result of actual system performance, consumer misperceptions of that performance or the importance consumers place on various perceptions. Finally, the model provides an understanding of transportation consumers which can serve as the basis for developing a broad range of strategies to influence consumer mode choice decisions. Using this model strategies which focus on directly altering consumer perceptions or choice may be considered in addition to traditional service modification strategies. For these reasons the model in Figure 1 was used as a basis for the study reported here.

The major objective of our research is to examine the relationships between consumers' mode perceptions, preference, and choice (i.e. the aggregation process) as a basis for developing strategies to modify consumers'

choice. We also examine the adequacy of the model in ordering the relationships between system characteristics, perceptions, preference and choice. However, we do not address the abstraction process in this paper since insufficient data was collected to fully investigate it.

The Study

Context

The research was conducted in the city of Evanston, Illinois, in cooperation with the City Manager's Office. Evanston is a northern suburb of Chicago, with a population of approximately 80,000. The Evanston public transit system includes: a rapid transit system which serves Evanston and connects with the Chicago rapid transit system, access to the Chicago Northwestern Railroad which runs directly from the northern suburbs to downtown Chicago, and a local bus service.

The transit problems of the City of Evanston are typical of many suburban cities. Poor public transit service (i.e. variability in lead time and run time, lack of coordination between modes, and a system not tailored to the needs of the community) combined with relatively high auto ownership has resulted in significant excess capacity on the public transit system, especially during off-peak hours. As a result of low ridership and high fixed operating costs, an annual subsidy of \$300,000 is required to maintain the transit system. Therefore the City of Evanston provides a good context for research designed to understand consumers' travel behavior and evaluate strategies for increasing public transportation ridership.

Method

Our approach to gaining an understanding of consumer travel behavior entailed developing an instrument for collecting individual data on the variables in the model in Figure 1 and then using this data to test the relationships between these variables and to generate and evaluate strategies designed to alter consumers' behavior. In this section the development of the research instrument and the data collection process are described.

The primary research instrument used was a set of mail questionnaires. Questionnaires were selected for the data collection because they provided the most efficient means of collecting quantitative data on the major variables in the model for a large cross section of the population.¹ The questionnaires were administered by mail because this method allowed us to reach a broader cross section of the population than alternative methods (i.e. telephone or personal interview) and it allowed consumers to respond to the questionnaires, which were quite lengthy, at their leisure.

The development of the questionnaires was a complex, lengthy process. First, input was obtained from three major sources; a usage audit, focus group interviews

¹However, the use of a questionnaire, especially a mail questionnaire made it impossible to assess actual system characteristics for the respondents at the same point in time as other variables were measured. Instead only perceptions of system characteristics were obtained. A separate study was conducted to explore the relationship between actual and perceived system performance. This study found no significant differences between actual and perceived measures of these factors. Therefore, it appears that these measures can be used interchangeably in this context.

and a review of the literature. The usage audit served to identify the most frequently used modes of transportation for various trip purposes within the community. These frequently used modes were then included in the questionnaires for consumer evaluation (space and time constraints prohibited requesting individuals to evaluate all modes). The focus group interviews and the literature review served to : 1) generate a list of transportation service attributes important to consumers to be used in evaluating modes in the final questionnaire and 2) provide "semantics" so that the questionnaire could be phrased in the language of the consumer. Since these inputs indicated that there were differences in the transportation alternatives available and the relevance of various transportation attributes for different types of trips, three different versions of the questionnaire were developed for each of three major types of trips: trips to work or school, non-work trips to the CBD, and non-work trips to areas in the city other than the CBD.

On the basis of the focus groups and literature review, scales measuring attributes of transportation services important to consumers were developed and evaluated. This entailed generating an exhaustive list of attributes and then reducing this list to a manageable set of critical attributes (21-25 scales per mode) on the basis of the results of two pretests in which individuals in the community used the attributes to evaluate several modes and also indicated their mode preference and choice. (For more details on the scale development process see Hauser, Tybout, and Koppelman, 1977).

Once the scales were developed, the remaining sections of the questionnaires were drafted and a pretest of the entire questionnaire was conducted. Results of this pretest were used to uncover and correct any problems or omissions in the questionnaire. In addition, the completed pretest questionnaires were used in a pre-analysis. This process entailed performing a full-scale statistical analysis on these responses to ensure that the questionnaire provided all the data necessary for the analysis planned on the final questionnaire. A few missing items were discovered and subsequently added, but the most important result of preanalysis was the recognition that mode-specific attributes, popular in the transportation literature, are not compatible with perceptual models unless certain key steps are taken. These steps require that either there be parallel generic attributes for all modes or if attribute has no parallel for a mode (i.e. walk access for the walk mode) or there be a clear extreme direction for that mode (i.e. walk access is zero). Based on the pretest and pre-analysis the final questionnaires were developed, printed and mailed to a random sample of the target population (1900 work/school trip, 1900 nonwork trip to the CBD and 950 nonwork trip to non-CBD destinations questionnaires were sent.) A follow up post-card was sent to every respondent seven days after the questionnaire mailing, urging them to return the questionnaire.

A summary of the sections of the questionnaire used in the analysis reported here is given below.

- (1) system characteristics. Respondents gave estimates of: travel time, broken down by access, wait, and on-vehicle time; bus frequency during rush and non-rush hours; the distance to the nearest bus stop; and bus seat availability. They also provided data which was used to compute auto availability (i.e. they reported the number of drivers and autos in their household).
- (2) perceptions. Respondents evaluated each of the three frequently used modes (car, bus, and walk) by responding to 21-25 statements about mode attributes on a 5 point, strongly agree to strongly disagree Likert scale. Respondents also expressed their feelings about each mode in terms of their affect, personal

normative beliefs, social normative beliefs and level of commitment by responding to 6-9 statements regarding these factors on a 5 point Likert scale from strongly agree to strongly disagree.²

- (3) preference. Respondents rank ordered the three modes - bus, auto and walk - in terms of their preference.
- (4) choice. Respondents indicated the mode which they had used for their most recent nonwork trip to the CBD and also estimated the frequency with which they had used each of the available modes for similar trips in the past two months.
- (5) consumer and situational differences. Participants responded to a battery of demographic questions (i.e. age, income, education etc.) and also described characteristics of their most recent trip to downtown Evanston (i.e. purpose, time of day etc.).

Results³

Respondents. Forty-one percent (782 out of 1900) of the individuals who received the nonwork trip to downtown Evanston returned it. Five hundred of these responses were selected for analysis on the basis of completeness of response to the attribute ratings and preference rankings questions. Comparison of the demographic characteristics of this sample with 1970 census data indicated that it was reasonably representative of the Evanston population, although some distinctions did exist (see Koppelman, Hauser and Tybout, 1977, for details on data screening and sample representativeness).

Description of System Characteristics. Although consumers' perceptions of a variety of system characteristics (i.e. travel time, frequency, bus seat availability, bus accessibility and auto availability) were measured, some of these variables require special coding before they can be used in the analysis. Therefore, only travel time and car availability are discussed here.

Walk travel time for trips to downtown Evanston was perceived to range between 1-5 minutes and 86-90 minutes, while vehicular travel time for similar trips ranged from 1-5 minutes to 26-30 minutes for car and 1-5 minutes to 56-60 minutes for bus. Car availability ranged between 0.0 auto/driver to 3.0 autos/driver, however 97% had one or less autos/driver.

Description of Perceptions of Mode Attributes. The questionnaire measured consumers' perceptions of three modes - bus, walk, and car passenger or driver - on 25 attributes. Respondents' evaluations of the modes on these attributes are summarized in Table 1.

²Historically, transportation researchers have only focused on one psychological dimension, beliefs about attributes of the object (e.g. perception of mode convenience, comfort etc.). Many other psychological variables have been demonstrated to influence behavior, including affect (an individual's liking-disliking of an object, see Ostrom, 1969), personal normative beliefs (an individual's perception of what he ought to do, see Schwartz and Tessler, 1972), social normative beliefs (an individual's perception of what others want him to do, see Fishbein, 1972), and level of commitment, (how easily is the individual's intended behavior influenced by unanticipated events, see Wicker, 1971). These perceptions have been included in our questionnaire so that their contribution to the explanation of transportation behavior can be examined.

³Analysis of this data set is still in process, therefore these results are preliminary.

TABLE 1
AVERAGE STANDARDIZED ATTRIBUTE RATINGS

	BUS	WALK	CAR
ON TIME	-0.12	.08	1.13
NO TRIP SCHEDULING NECESSARY	-.78	-.48	-.07
RELAXING	.85	1.00	.78
CORRECT TEMPERATURE	.31	-.15	.80
NO WORRY OF ASSAULT	.76	.38	1.06
CAN COME AND GO AS I WISH	-.47	.67	.83
INEXPENSIVE	.56	1.10	-.59
ERRANDS TAKE LITTLE TIME	-.28	-.36	.81
NO WORRY ABOUT INJURY	.98	.68	.71
KNOW HOW TO GET AROUND	.73	1.04	.99
LITTLE EFFORT INVOLVED	.26	-.21	.64
AVAILABLE WHEN NEEDED	-.20	.91	.56
NOT MADE UNCOMFORTABLE BY OTHERS	.91	1.01	.90
NO PROBLEMS IN BAD WEATHER	.01	-.73	.30
PLEASANT DRIVERS OR OTHER PERSONNEL	.43	.43	.41
GET TO DESTINATION QUICKLY	-.09	-.50	.84
PROTECTED FROM SMOKING	.09	.65	.75
SAFE AT NIGHT	-.02	-.51	.68
NOT ANNOYED BY OTHERS	.74	.81	.57
NO LONG WAITS	-.30	.77	.75
EASILY CARRY PACKAGES	-.19	-.57	1.03
EASY TO TRAVEL WITH SMALL CHILDREN	-.01	-.37	.75
NOT TIRING	.44	-.30	.82
EASY GETTING IN AND OUT	.56	1.27	.82
EASY WALK ACCESS	.79	1.27	.96

*THE RATINGS WHICH APPEAR IN THIS TABLE WERE STANDARDIZED BY INDIVIDUALS ACROSS STIMULI AND SCALES TO REMOVE ANY TENDENCY OF AN INDIVIDUAL TO USE ONLY PART OF THE RANGE IN THE SCALE. IN ADDITION, ALL NEGATIVELY WORDED SCALES WERE MATHEMATICALLY RESERVED SO THAT HIGHER NUMERICAL VALUES IMPLY BETTER RATINGS.

An examination of the ratings in Table 2 indicates that, in general, car is perceived favorably and outscores bus and walk. However, on cost ("inexpensive") and driver stress attributes ("fear of injury", "annoyed by others") car fares less well. In contrast, bus is relatively poorly perceived. It receives high ratings only on cost and stress related attributes and is viewed negatively in terms of service attributes. Walk is rated highly in terms of attributes measuring cost, service availability and environment, but is also seen as time consuming and requiring considerable effort to use.

Factor analysis was used to reduce these 25 transportation service attributes to a smaller set of underlying dimensions. This was done for two reasons: 1) consumers do not actually process information about each of the 25 attributes when making an evaluation or choice regarding transportation alternatives (see Bruner, Goodnow and Austin, 1959). Instead they reduce the information to a smaller more manageable set of factors that capture the essence of the larger set. Thus, a simpler perceptual structure more closely approximates consumers' utilization of perceptual information in decision making. 2) This simpler structure helps managers and analysts better understand consumer processes so that they can formulate strategies to affect the most crucial components of consumer response. And, 3) factor analysis enables the analyst to make dimensions orthogonal, thus reducing multicollinearity and leading to stable coefficients when the dimensions are used in preference and choice models.

Factor analysis of the attribute ratings was undertaken for two through six dimensions using common factor analysis with iterations and varimax rotation. The solutions for the various dimensions were compared on the basis of interpretability, explanatory power and accuracy in predicting preference. On the basis of this analysis the three dimension solution was chosen as the best one.

The factor loadings for the three dimension solution appear in Table 2 and mode perceptions on these dimensions are graphed in Figure 2. The three factors - which have been labeled general service and safety, convenience and accessibility and psychological comfort - account for 45% of the variance in the original attribute ratings. This is consistent with previous studies of this nature (see

Hauser and Urban, 1977). As Figure 2 demonstrates, car is perceived most favorably on the general service dimension but is viewed less favorably than alternative modes on the convenience/accessibility and psychological comfort dimensions. Bus is viewed poorly on the general service and convenience/accessibility dimension, but receives the most favorable evaluation on the psychological comfort (i.e. freedom from hassles) dimension. Finally, walk scores extremely poorly on the general service dimension, well on the convenience/accessibility dimension, and moderately well on the psychological comfort dimension.

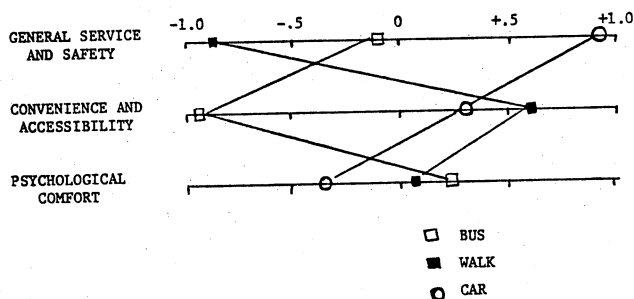
TABLE 2
FACTOR LOADINGS FOR ATTRIBUTE RATINGS

	FACTOR 1	FACTOR 2	FACTOR 3
ON TIME	.57	.40	-.08
NO TRIP SCHEDULING NECESSARY	.26	.27	-.27
RELAXING	.48	.14	.26
CORRECT TEMPERATURE	.58	.01	.11
NO WORRY OF ASSAULT	.48	.00	.30
CAN COME AND GO AS I WISH	.25	.68	-.04
ERRANDS TAKE LITTLE TIME	.69	.29	-.08
NO WORRY ABOUT INJURY	.18	-.07	.47
KNOW HOW TO GET AROUND	.09	.33	.20
LITTLE EFFORT INVOLVED	.69	.09	.21
AVAILABLE WHEN NEEDED	.02	.67	.09
NOT MADE UNCOMFORTABLE BY OTHERS	.06	.22	.54
NO PROBLEMS IN BAD WEATHER	.62	-.03	.14
PLEASANT DRIVERS OR OTHER PERSONNEL	.06	.09	.33
GET TO DESTINATION QUICKLY	.77	.16	-.03
PROTECTED FROM SMOKING	.12	.38	.04
SAFE AT NIGHT	.62	.00	.10
NOT ANNOYED BY OTHERS	.04	.12	.51
NO LONG WAITS	.16	.64	-.03
EASILY CARRY PACKAGES	.71	.13	-.08
EASY TO TRAVEL WITH SMALL CHILDREN	.59	.06	-.08
NOT TIRING	.77	-.00	.19
EASY GETTING IN AND OUT	-.15	.51	.29
EASY WALK ACCESS	-.12	.48	.28

FACTOR INTERPRETATION

FACTOR 1 - GENERAL SERVICE AND SAFETY
FACTOR 2 - CONVENIENCE AND ACCESSIBILITY
FACTOR 3 - PSYCHOLOGICAL COMFORT

FIGURE 2
MODE PERCEPTIONS IN THREE FACTOR SPACE



Description of Feelings About Modes (Perceptions of Factors other than Mode Attributes). In an effort to determine whether psychological or perceptual factors other than evaluations of mode attributes influence transportation preference and choice, a variety of non-attribute perceptions of the three modes were measured (i.e. affect, personal normative beliefs, social normative beliefs, extraneous events). These measures were then factor analyzed to develop an aggregate measure of feelings toward each mode.⁴ The factor loadings for the

⁴This approach was taken because each of the original variables was measured by only a few questions per mode and therefore these measures were likely to be unstable if used separately. However, when the variables are combined for each mode they provide a fairly reliable index of a more general feeling toward the mode.

three resulting factors - a car feelings factor, bus feelings factor and walk feelings factor - are presented in Table 3. These factors account for 39% of the variance in the original set of questions.

TABLE 3
FACTOR LOADINGS FOR FEELINGS

	WALK FEELINGS	BUS FEELINGS	CAR FEELINGS
DIFFERENT FROM BUS RIDERS	.05	-.28	.01
ENJOY TRAVEL BY CAR	-.23	-.01	-.55
ENJOY TRAVEL BY BUS	-.12	.71	-.13
ENJOY TRAVEL BY FOOT	.83	.07	-.04
DEPRESSING TO TRAVEL BY CAR	.11	-.06	.76
DEPRESSING TO TRAVEL BY BUS	.04	-.53	.41
DEPRESSING TO TRAVEL BY FOOT	-.67	-.11	.31
PEERS SURPRISED IF RIDE BUS REGULARLY	-.08	-.52	-.06
OUGHT TO TRAVEL BY CAR	-.45	-.19	-.38
OUGHT TO TRAVEL BY BUS	.01	.52	.16
OUGHT TO TRAVEL BY FOOT	.75	-.01	.10
PEERS SURPRISED IF DROVE CAR REGULARLY	.18	.16	.32
IF WEATHER BAD, FEWER CAR TRIPS	-.07	.10	.31
IF WEATHER BAD, FEWER BUS TRIPS	-.14	-.34	.07
IF WEATHER BAD, FEWER WALK TRIPS	-.25	-.03	-.05
IF GASOLINE PRICE DOUBLED, MORE CAR TRIPS	-.30	-.21	-.14
IF GASOLINE PRICE DOUBLED, MORE WALK TRIPS	.69	.06	.16
IF GASOLINE PRICE DOUBLED, MORE CAR POOL TRIPS	.12	.18	.03
IF GASOLINE PRICE DOUBLED, FEWER CAR ALONE TRIPS	.15	.26	.09
PEERS SURPRISED IF WALKED ALOT	-.66	-.11	-.11
IF BUS FARES LOWER, MORE TRIPS BY BUS	.09	.61	.35
IF BUS FARES LOWER, FEWER TRIPS BY CAR	.06	.59	.34
IF BUS RAN MORE OFTEN, MORE BUS TRIPS	.21	.40	.27
WOULD TRAVEL BY CAR REGARDLESS OF COST	-.40	-.39	-.40
WOULD TRAVEL BY BUS EVEN IF LONG WALK	.12	.50	.11
IF PARKING COST DOUBLED WOULD WALK	.44	.04	.21
WILLING TO CAR POOL SOME TRIPS	.06	.09	.00

Preference and Choice. First preference was clearly dominated by car (71% stated this as their first preference), while bus did well in terms of second preference (58% stated bus was their second preference). Consistent with the preference ratings, 66% stated that they had chosen car for their most recent trip. Furthermore, cross-tabulation of first preference and choice data indicated that the vast majority of respondents (76%) chose their most preferred mode. However, it is interesting to note that a significant number of individuals (24%) did not choose their most preferred mode, perhaps due to situational constraints such as availability. This highlights the importance of consumers' second preferences and suggests an area of opportunity for public transportation.

Relationship Between Model Variables. Our model of consumer travel behavior states that the impact of system characteristics on preference and choice is mediated by consumer perceptions. Therefore, system characteristics should be more highly correlated with perceptions of mode attributes than with preference. Examination of Table 4 indicates that this is the case for the travel time variables. Walk travel time is most highly correlated with perceptions of general service ($r = -.63$) while vehicular travel time is most highly correlated with perceptions of convenience/accessibility ($r = -.41$). However, auto availability (APD) is highly correlated with not only perceptions of general service, but also preference and choice. This is probably partially attributable to the fact that none of the attribute scales directly measured mode availability. However, auto availability may also have a high correlation with choice because it operates as a situational constraint rather than a system characteristic. If this is the case auto availability is probably best used as a segmentation variable rather than as an independent variable in mod-

els predicting perception, preference or choice.

TABLE 4
CORRELATION MATRIX

	SYSTEM CHARACTERISTICS			ATTRIBUTE PERCEPTIONS			FEELINGS TOWARDS MODES			P
	WTT	VTT	APD	GS	C/A	PC	BF	WF	CF	
WALK TRAVEL TIME (WTT)										
VEHICULAR TRAVEL TIME (VTT)	-.33									
AUTOS/DRIVER (APD)	-.29	.01*								
GENERAL SER- VICE FACTOR SCORE (GS)	-.63	.10	.50							
CONVENIENCE/ ACCESSIBILITY FACTOR SCORE (C/A)	.30	-.41	.25	.04						
PSYCHOLOGICAL COMFORT FACTOR SCORE (PC)	-.01*	.01*	-.20	.03*	.04*					
BUS FEELINGS FACTOR SCORE (BF)	.00*	-.14	.00*	.17	.13	.14				
WALK FEELINGS FACTOR SCORE (WF)	-.20	.00	.00*	.31	.10	.11	.00*			
CAR FEELINGS FACTOR SCORE (CF)	.00*	.05	-.06	-.07	-.09	-.13	.00*	.00*		
PREFERENCE (P)	-.28	-.04	.50	.56	.31	-.05	.13	.22	-.11	
CHOICE (C)	-.30	-.02	.54	.52	.31	-.09	.13	.17	-.11	.66

*NONSIGNIFICANT AT .05 LEVEL

The perceptual variables (cognitive dimensions and feelings) are viewed as the determinants of preference. Therefore, these variables should be relatively independent and highly correlated with preference. In general these conditions are met. The intercorrelations between these variables are low and they are all significantly related to preference.

Finally, preference and situational constraints are viewed as the determinants of choice. Thus, choice should be highly correlated with these variables. Consistent with this expectation, Table 4 indicates that choice is most highly correlated with preference ($r = .66$) and next most highly correlated with a situational constraint - auto availability ($r = .54$). Therefore, we conclude that examination of the correlation matrix provides support for our model of consumer transportation behavior. Next, we will examine the aggregation process posited by this model in more detail.

The Aggregation Process - Preference Models. Preference models based on consumer perceptions were developed. First preference logit was used to statistically estimate the importance weights for the variables in the models. These models, which are presented in Table 5, are summarized below.

In Model 1 the three cognitive dimensions (factor scores) are used to predict preference. The importance weights for these dimensions in this model follow the same pattern as their identification in factor space. General service is the most important variable. Convenience and accessibility is next most important. And, psychological comfort is least important.

In Model 2, the factor scores for respondents' feelings about each mode are added to Model 2. The addition of these variables improves the prediction and explanation of preference by a small but statistically significant amounts ($p < .01$). This finding supports our belief

that perceptual variables other than evaluations of mode attributes influence preference and suggests that efforts to alter preference should consider these variables.

TABLE 5
PREFERENCE MODELS

VARIABLE NAME	FIRST PREFERENCE MODELS	
	(1)	(2)
GENERAL SERVICE FACTOR SCORE	.57	.40 (.68)*
CONVENIENCE/ACCESSIBILITY FACTOR SCORE	.32	.18 (.31)
PSYCHOLOGICAL COMFORT FACTOR SCORE	.11	.01 (.02)
CAR FEELINGS FACTOR SCORE		.08
BUS FEELINGS FACTOR SCORE		.15
WALK FEELINGS FACTOR SCORE		.19
PERCENT PREDICTED INFORMATION	78.1% 53.3%	80.8% 56%

*WEIGHTS IN BRACKETS ARE STANDARDIZED IMPORTANCE WEIGHTS FOR THE COGNITIVE DIMENSIONS ALONE IN THIS MODE.

In general, the models presented in Table 5 do a good job of predicting first preference. The percent of first preferences correctly predicted by these models are all significantly higher than the percent which would be correctly predicted using a market share model (54.7%) or an equally likely model (33.3%). The information measure reported gives the percent of uncertainty (entropy) explained. This measure is an information theoretic interpretation (Hauser, 1977) of a pseudo R² measure (McFadden, 1970). Similar models were developed to predict rank preference using a rank logit model and the rank ordering of variables and interpretation were the same.

Choice Models. Since situational variables as well as preference affect choice a related set of models were developed to predict choice. These models, which were based on the multinomial revealed preference logit formulation (McFadden, 1970) are summarized in Table 6.

Model 1 in Table 6 uses only factor scores for the three cognitive dimensions to predict mode choice. In this model the importance weights for the dimensions were determined in the choice model. This procedure for determining the importance weights was compared to the alternative procedure of using importance weights established in the first preference Model 1. The two models yielded highly similar results. The only difference was the lowered importance of psychological comfort in the choice model. (Future papers will explore this issue further). Since most interpretations were similar for the two procedures and prediction was improved significantly (.05 level) when the importance weights were determined in the choice model, this procedure was used in the models reported here.

In Model 2 the factor scores for feelings about the three modes are added to Model 1. The addition of these three factors improves the prediction and explanation of choice behavior by a small but significant (.01 level) amount.

TABLE 6
MODE CHOICE MODELS

VARIABLE NAME	MODEL	
	(1)	(2)
GENERAL SERVICE FACTOR SCORE	.59	.35
CONVENIENCE/ACCESSIBILITY FACTOR SCORE	.39	.22
PSYCHOLOGICAL COMFORT FACTOR SCORE	.02*	-.05*
CAR FEELINGS FACTOR SCORE		.10
BUS FEELINGS FACTOR SCORE		.18
WALK FEELINGS FACTOR SCORE		.10
PERCENT PREDICTED INFORMATION	77.5% 48.1%	78.9% 50.0%

*ALL VARIABLES EXCEPT THOSE STARRED ARE SIGNIFICANT AT THE .05 LEVEL

In summary, the choice models in Table 6 generally do an excellent job of predicting mode choice for respondents most recent trip to downtown Evanston. All predictions are substantially better than market share or equally likely models would allow. Models predicting respondents' reported frequency of choice over the last two months were also developed. In general these models were similar to those shown in Table 6.

Segmentation Using Consumer and Situational Characteristics. As depicted in the model (Figure 1), consumer and situational characteristics may influence mode preference and choice by influencing the aggregation process or by influencing choice directly. When these factors have a substantial impact on this process, a clearer understanding of consumer transportation behavior may be obtained by segmenting on these variables and developing separate choice and preference models for each segment.

Two approaches to segmentation may be employed. One approach entails identifying segments in the population on the basis of differences in the descriptive characteristics of individuals (i.e. demographics) or situations (i.e. trip characteristics). An alternative approach involves identifying segments that are behaviorally homogenous and distinct (i.e. segments that respond to transportation alternatives similarly and differently than other segments). In an effort to increase our understanding of consumer transportation behavior we employed both of these approaches to segmentation. First, we examined the impact of segmenting on each of four individual and situational descriptive characteristics - age, education, knowledge of the bus system and auto availability - on the prediction and explanation of mode choice.⁵ A Chi square test was used to assess the significance of each basis of segmentation. Only two of the four bases, age and auto availability, were significant. Choice models developed for the subgroups of these significant bases of segmentation are reported in Table 7.

TABLE 7
SEGMENTED CHOICE MODELS

VARIABLE NAME	AUTO AVAILABILITY		AGE		
	LOW ¹	HIGH ²	UNDER 30	30-59	60 or older
	GENERAL SERVICE FACTOR SCORE	.24	.42	.27	.61
CONVENIENCE/ACCESSIBILITY FACTOR SCORE	.20	.21	.27	.22	.19
PSYCHOLOGICAL COMFORT FACTOR SCORE	.02*	-.08	-.07*	-.05*	-.02*
CAR FEELINGS FACTOR SCORE	.15	.07*	.17	-.05*	-.07*
BUS FEELINGS FACTOR SCORE	.24	.15	.09*	.06*	.31
WALK FEELINGS FACTOR SCORE	.15	.06*	.13	.01	.20*
% CORRECTLY PREDICTED BY SEGMENT	74.3	80.2	80.0	85.1	77.0
INFORMATION BY SEGMENT	51.0	54.0	40.0	63.0	54.0
X ² FOR SEGMENTATION	23.5, p<.05		40.6, p<.01		
TOTAL % CORRECTLY PREDICTED	78.9%		81.9%		
TOTAL INFORMATION	53.3%		56.1%		

¹ LESS THAN .5 AUTOS/DRIVER
² .5 OR MORE AUTOS/DRIVER

*NONSIGNIFICANT AT THE .05 LEVEL

The choice models for the auto availability subgroups indicate that those with low auto availability place less

⁵ Since preference and choice models were highly similar only choice models were examined in the segmentation analysis. Tests of other descriptive bases of segmentation will be conducted in the future.

emphasis on the general services offered by alternative modes and are influenced more by feelings about bus and walk than those with high auto availability. Similarly, the choice models for the age subgroups indicate that individuals under thirty are primarily influenced by the general service and the convenience/accessibility of modes and their negative feelings about car. In contrast the mode choice among middle-aged individuals appears to be primarily determined by general service. The elderly, although significantly influenced by mode attributes (i.e. general service and convenience/accessibility), are most heavily influenced by their feelings about the bus in making mode choice decisions.

An attempt was also made to segment on the basis of individuals' response to transportation alternatives. To identify individuals with similar behavioral responses, a cluster analysis was performed on the factor scores for feelings toward the modes. Using this procedure four segments were identified. These segments represented different reactions to the available modes; one was open-minded (i.e. positive about all modes), one was anti-walk (i.e. negative about walk, neutral about other modes), one was anti-car (i.e. negative about car, positive about other modes), and one was anti-bus (i.e. negative about bus, positive about other modes). Furthermore these segments represented different types of people (e.g. anti-car tended to be male university students with relatively low income etc.). However, this basis for segmentation was not significant in terms of improving the prediction and explanation of mode choice, therefore the models for these subgroups are not reported here.

In summary, the segmentation analysis demonstrates that a significantly better explanation of consumer travel behavior may be obtained by breaking the population down into distinct subgroups. This increased understanding can be helpful in guiding strategy development.

Discussion and Plans for Future Research

In general, the research results support our model of consumer transportation behavior. However, a more stringent test of the model requires the manipulation of model variables and observation of their effects. This test is necessary not only from a theoretical perspective (i.e. to test the causality of hypothesized relationships between variables), but also from a practical perspective. The model is only of value to transportation managers and planners if it can help them generate and select effective strategies for influencing consumer mode choice. Therefore, in this section, the research results will be used to generate strategies for altering consumers' mode choice and future research to evaluate these strategies will be outlined.

In accord with our model, five basic types of strategies for influencing mode choice may be generated. These strategies, each focusing on a different model variable, are summarized below.

- (1) Modification of System Characteristics. Strategies designed to modify system characteristics (i.e. product strategies) are appropriate when the system does not meet consumers' needs on some dimension and this dimension is significantly related to perceptions which influence preference and choice. System strategies may range from reducing bus fares or increasing bus frequency to introducing a new paratransit mode.
- (2) Modification of Consumer Perceptions. Strategies designed to modify consumer perceptions directly (i.e. not by modifying system characteristics) are appropriate when consumers are either uninformed or misinformed about the system. Here the task is to provide consumers with accurate information. In addition, modification of consumer perceptions may also be ap-

propriate when consumers have accurate information about the system but interpret that information negatively (e.g. they know that the bus runs every half hour and interpret that as poor service). When this occurs persuasion may be employed in an effort to alter the individuals' interpretation of the information.

- (3) Modification of Consumer Preference. Strategies designed to modify consumer preference may be employed when consumers' perceptions are accurate but low importance is placed on perceptions of public transportation which are positive and high importance is placed on those which are negative. In this situation the task is one of changing the importance weights so that dimensions on which public transportation performs well receive greater emphasis. Persuasive appeals may be used to do this.
- (4) Modification of Situational Constraints. Strategies designed to modify situational constraints are appropriate when these factors have a significant impact on mode choice (i.e. auto availability, parking availability etc. may influence mode choice). This approach entails manipulation of situational factors so that incentives or disincentives for particular mode choices result (e.g. restricting parking in downtown areas may discourage car trips). These strategies tend to be perceived as more coercive than other types of strategies and they are often difficult to implement since legislation and regulation changes may be required. Thus, constraint modification strategies are typically only used when other strategies have failed.
- (5) Modification of Behavior. Strategies designed to modify individuals' choice behavior directly are appropriate when consumers are reluctant to try a particular service but there is reason to believe that if they tried it they would like it. Promotional strategies such as free rides, discount coupons etc. may be used to encourage trial.

Now, let us use the research findings to determine what strategies are likely to increase public transit ridership in the Evanston community. First, we will examine strategies for the entire community. Then, we will discuss strategies tailored to the behavior of subgroups in the community.

The research results indicate that perceptions of general service provided by alternative modes are the most important determinant of preference and choice for the community as a whole. Thus, one strategy for increasing public transportation ridership is to improve the perceptions of the general service public transportation provides. To do this we must first examine the specific attributes which make up the general service dimension (Table 2). Attributes loading heavily on the general service factor are "on time" "correct temperature" "errands take little time", "little effort involved", "get to destination quickly", "no problems in bad weather", "easy to carry packages", "safe at night", "easy to travel with small children", and "not tiring". Of these attributes bus scores very poorly on the six that are underlined. Next, we must attempt to determine whether the low evaluation of bus on these dimensions is the result of poor system performance or a misperception of system performance by consumers.

Evidence from a separate study in which system performance was monitored indicates that, in general, the system runs on schedule (see Bernstein, Knall and Lindauer, 1977). Therefore it appears that the low "on time" rating for bus is not the result of poor system performance but rather it stems from consumers' misperceptions or lack of information regarding the system. This in-

terpretation is also supported by the finding that consumers generally could not correctly answer questions regarding the bus schedule. As a result, one strategy for improving bus general service would be to inform consumers about the good "on time" performance of the bus and educate them about the bus schedule.

In contrast, evidence from the same study of system performance suggests that consumers accurately perceive the bus travel time from their homes to downtown Evanston. Thus, the negative evaluation of bus on quickness related attributes is not due to a misperception but instead is the result of a negative interpretation of an accurate perception. Two alternative strategies for improving the "quickness" evaluation of bus and thereby improving the general service rating are appropriate: (1) modify the service (e.g., implement express buses) and (2) use persuasive communications to convince consumers that they should reinterpret the actual quickness more favorably.

Similarly, examination of the bus system performance indicates that consumers' negative perceptions of the ease of carrying packages and traveling with children on the bus are probably well founded. Perceptions of ease of carrying packages could be made more positive by altering the service to include special package racks, etc. Strategies which would increase the ease of traveling with children on the bus are not readily apparent.

Finally, consumers' poor rating of bus safety at night has some basis in reality (i.e. rapes and muggings have occurred in the past), however consumers may be overreacting to one or two isolated incidents that have been highly publicized. A two-pronged approach to improving evaluations to bus safety may be appropriate: (1) change bus system to improve safety (e.g. add guards, improve lighting etc.) and (2) inform public that actual crime rate is quite low.

The above strategies for increasing public transit ridership are all based on improving consumers' evaluation of bus on the general service dimension. An alternative approach would be to use persuasive communications to increase the relative importance of the psychological comfort dimension where bus already outperforms car and walk. This might entail stressing the importance of getting to one's destination without being hassled (e.g. Greyhound's "leave the driving to us"). Or, emphasis could be placed on improving consumers' feelings about the bus and the importance of these feelings. The bus feeling factor is a function of liking for the bus, believing one ought to ride the bus, and believing significant others (e.g. family, friends) would approve of one riding the bus. Thus, strategies might focus on making the bus ride more pleasant (e.g. attractive colors, music, etc.), emphasizing one's obligation to ride the bus to conserve energy, reduce environmental pollution and please one's friends, family, coworkers, etc. Finally, once changes in the bus system have been made and it appears that the service is one that meets the needs of the consumer, strategies to alter behavior (i.e. encourage trial), such as coupons, free rides, etc. may be appropriate to help alter consumer perceptions.

The same procedure used to develop strategies for the community as a whole can be used to develop strategies for subgroups in the population. Table 7 indicates that individuals who have low auto availability place less emphasis on general service, where bus performs poorly, and more emphasis on feelings about the bus than those with high auto availability. This suggests that situational constraints which reduce auto availability will have a positive impact on bus ridership. In addition, the data provide guidelines for the development

of specific strategies to reach each of these segments (i.e. focus on general service strategies for those with high auto availability and use strategies designed to improve bus feelings and general service strategies for those with low auto availability).

The choice models for the age segments also suggest using different strategies to increase public transportation ridership within age group. Members of the youth segment are strongly influenced by their feelings about the car, which tend to be negative. Thus, strategies to increase public transit ridership among the young could make their negative feelings about car salient and present bus as a more attractive alternative. The middle-aged segment is heavily influenced by perceptions of general service, therefore the strategies for improving general service discussed earlier are appropriate for increasing public transit ridership among this group. In contrast, the elderly's mode choice is largely influenced by their feelings about the bus, thus strategies for enhancing bus feelings discussed earlier are particularly appropriate for this segment. Furthermore general service and convenience/accessibility are of equal importance for the elderly and young, therefore strategies effecting either of these dimensions are likely to have a similar impact.

We have illustrated how the model can be used to generate strategies for influencing mode choice. The next step in our research program will be an evaluation of these strategies. This evaluation will entail comparing the estimated costs of implementing each strategy (i.e. materials/equipment personnel time etc.) with its estimated benefits (i.e. predicted impact on ridership, subsidies, etc.). (An on-line computer package is currently being developed for this purpose). In addition, strategies will be evaluated in terms of their feasibility (i.e. whether there are any legal constraints which would affect their implementation such as labor union regulations). On the basis of this evaluation a set of strategies will be implemented and their effects will be monitored to test the accuracy of our model of consumer behavior. For example, the model predicts that changes in general service will have the greatest impact on middle aged individuals and will have less impact on the young and the elderly. Furthermore, the model predicts that strategies to reduce travel time will have a greater impact on perceptions of general service than ones which increase availability (i.e. extend hours of operation). Implementation of two strategies, one which reduces travel time and one which extends the hours of service and monitoring the impact of these strategies through direct measurement of demand (i.e. ridership counts, fare box revenues etc.), and responses to consumer surveys (i.e. surveys measuring consumers' awareness, perceptions, preference and choice in response to these changes) would allow testing these model based predictions.

References

- M.E. Ben-Akiva, Structure of Passenger-Travel Demand Model, unpublished Ph.D. thesis, (Cambridge: M.I.T., 1973).
- Karen Bernstein, Michael Knall and Nancy Lindauer, "A Study to Evaluate Differences in Perception and System Performance for the Evanston Bus System" unpublished paper (1977).
- J.S. Bruner, J.J. Goodnow, and G.A. Austin, A Study of Thinking, (New York: Wiley and Sons, 1959).
- E. Brunswik, The Conceptual Framework of Psychology, (Chicago: University of Chicago Press, 1952).

Charles River Associates, Inc., "A Disaggregate Behavioral Model of Urban Travel Demand," (Federal Highway Administration, U.S. Department of Transportation, Washington D.C., 1972).

Martin Fishbein, "The Search for Attitudinal Behavioral Consistency," in Joel Cohen (Ed.), Behavioral Science Foundations of Consumer Behavior (New York: The Free Press, 1972).

Paul Green and Yorum Wind, "New Way to Measure Consumer's Judgements," Harvard Business Review, (July-August, 1975).

Kenneth R. Hammond, "Inductive Knowing" in Joseph R. Royce and William W. Rozeboom eds. The Psychology of Knowing, (New York: Gordon and Breach, 1972), 285-343.

John R. Hauser, "Testing the Accuracy, Usefulness and Significance of Probabilistic Choice Models: An Information Theoretic Approach," Operations Research (Forthcoming, 1977)

John R. Hauser, Alice M. Tybout and Frank S. Koppelman, "Consumer Oriented Transportation Service Planning: The Development and Implementation of a Questionnaire to Determine Consumer Wants and Needs," Transportation Center Research Report No. 3 in the COTSP Series, Northwestern University, 1977).

John R. Hauser and Glenn L. Urban, "Normative Methodology for Modeling Consumer Response to Innovations," Operations Research, (July-August, 1977), 579-619.

Frank S. Koppelman, John R. Hauser and Alice M. Tybout, "Consumer Oriented Transportation Service Planning: Analysis of Perceptions, Preference and Choice," Transportation Center Research Report No. 5 in the COTSP Series, (Northwestern University, 1977).

Brian V. Martin and Frederick W. Memmott, III, Alexander J. Bone. Principals and Techniques of Predicting Future Demand for Urban Area Transportation (Cambridge: M.I.T. Press, June, 1961).

Daniel McFadden, "Conditional Logit Analysis of Quantitative Choice Behavior," in P. Zarembka, ed., Frontiers in Econometrics (New York: Academic Press, 1970), 105-142.

J.M. McLynn and T. Woronka, Passenger Demand and Modal Split Models, prepared for Northeast Corridor Transportation Project, United States Department of Transportation, Report No. NEC TP-230, Bethesda, Md., 1969.

Gregory C. Nicolaidis, "Quantification of the Comfort Variable," Transportation Research 9(1975), 55-66.

Thomas Ostrom, "The Relationship Between the Affective, Behavioral and Cognitive Components of Attitude," Journal of Experimental Social Psychology 5(1969), 12-30.

E.A. Pessemier, Product Management: Strategy and Organization (New York: Wiley/Hamilton, 1977).

R.E. Quandt and W. J. Banmol, "The Demand for Abstract Transportation Modes: Theory and Management," Journal of Regional Science 6(1966).

S. Schwartz and R. Tessler., "A Test of a Model for Reducing Measured Attitude-Behavior Discrepancies," Journal of Personality and Social Psychology 24(1972), 225-236.

Bruce D. Spear, "A Generalized Attribute Variable for Models of Mode Choice Behavior," a paper presented at the 55th Annual Meeting of the Transportation Research Board (January, 1976).

Systems Analysis and Research Corporation (SARC), Demand for Intercity Passenger Travel in the Washington Boston Corridor. Report to the U.S. Department of Commerce (1963).

Alan Wicker, "An Examination of the 'Other Variables' Explanation of Attitude-Behavior Inconsistency," Journal of Personality and Social Psychology 19(1971), 81-30.

FUNCTIONAL MEASUREMENT ANALYSIS OF SPATIAL AND TRAVEL BEHAVIOR

Jordan J. Louviere, University of Wyoming
 Irwin P. Levin, University of Iowa

Abstract

This paper overviews theory and methodology derived from Information Integration Theory or Functional Measurement and illustrates its application to several problems in modeling spatial choice and travel behavior. The applications involve both laboratory and field studies of human judgment: Consumer self-estimates of patronage on alternative public bus systems; consumer estimates of the attractiveness of alternative supermarkets; and consumer estimates of the attractiveness of alternative towns as places to live. All studies demonstrate highly significant monotonic relationships between modeled evaluations and corresponding real-world choice behavior.

Introduction

The purpose of this paper is to briefly overview the theory and methodology of Functional Measurement and illustrate its application through a brief discussion of several studies of spatial choice and travel behavior. Functional Measurement refers to a three-pronged approach to understanding and modeling human judgment and decision-making. The approach was developed by Norman Anderson and his associates (Anderson, 1970, 1976a, b, 1977; Shanteau and Troutman, 1975; Levin and Gray, 1976; Louviere, 1978; Louviere and Norman, 1977; Norman and Louviere, 1974; Norman, 1977). These three prongs are:

1. The establishment of interval scale measures for stimuli and response;
2. The establishment of the functional rule relating stimuli to response;
3. The establishment of the psychophysical rule relating physical and subjective measures of stimuli.

Functional Measurement Theory

Thus, the goal of the Functional Measurement approach is to develop mathematical expressions which describe the processes by which individuals and groups thereof process stimulus information in order to make judgments, decisions, choices, etc. The theory is concerned with multiple causation in that it postulates that individuals evaluate or make decisions about multi-attribute alternatives by means of simple algebraic "mental rules." In the present context, we would say that consumers of transportation products evaluate these products by means of simple algebraic combinations of attributes which the alternatives possess. As an example, we might postulate that consumers choose among or allocate their choices among alternative transportation modes for their trips to work by evaluating salient attributes of each mode, combining these separate evaluations into an overall mode evaluation in a manner which can be described algebraically, and selecting the mode with the highest evaluation most often. We might suppose, for example, that an individual weights each attribute separately and sums all the weighted attribute evaluations to assign an overall evaluation. This would lead to the following linear model(s):

$$E_i = \sum_j w_{ij} e_{ij} \quad (1)$$

$$E_i = \frac{\sum_j w_{ij} e_{ij}}{\sum_j w_{ij}} \quad (2)$$

depending upon whether the process is additive or averaging. E_i is the overall evaluation assigned to alternative i , w_{ij} is the weight assigned to each of the separate evaluations (e_{ij}) of the j attributes.

This has been a common assumption in much empirical work in marketing, geography, psychology and transportation (see Slovic and Lichtenstein, 1971; Slovic, Fischhoff, and Lichtenstein, 1977; and Wilkie and Pessemier, 1973 for a review). Unfortunately, when this assumption is put to the adequate test demanded for these models, as in the empirical work to be reported below, it has often been rejected. Rather, support has generally been given for one or the other of two alternative models:

$$E_i = \frac{\sum_{ij} w_{ij} e_{ij}}{\sum_{ij} w_{ij}} + k'' \quad (3)$$

$$E_i = \pi \prod_j w_{ij} e_{ij} + k' = k \pi e_{ij} + k' \quad (4)$$

Equation 3 is the expression for a differentially weighted averaging model in which the weight parameters (w_{ij}) are different for each different level (i) of each different attribute. Thus, there would be $i \times j$ weight values. The model is nonlinear because the denominator is different (non-constant) for each ij combination. Equation 4 is the multiplicative rule for the evaluations. Even though the attributes may have different weights, the weights are not mathematically distinguishable. Both equations 3 and 4 have intercept terms (k' , k'') to permit the existence of an arbitrary zero in the response scale (E_i), the requirement for an interval measure of evaluation.

There is an important conceptual difference between the models expressed in equations 3 and 4 and those expressed in equations 1 and 2. The linear models (equations 1 and 2) imply that the various attributes have independent effects on the overall evaluation of a multi-attribute system. By contrast, the nonlinear models (equations 3 and 4) imply that one or more attributes can modify the effect of others. Such modifying effects can be an important component of consumer decision processes. In an example to be described in more detail later, the effect of the sex of a potential carpool rider on the desirability of carpooling depended on whether or not the rider was a prior acquaintance of the respondent. Statistically, this difference corresponds to lack of interaction effects for the linear models and the discovery of systematic interaction effects for the nonlinear models.

Thus, the theory and methodology of functional measurement is concerned with the diagnosis and testing of these and other algebraic rules. In spirit, therefore, it is similar to conjoint measurement (Krantz and Tversky, 1971) or multi-attribute utility theory (Keeney and Raiffa, 1976) but is superior on methodological grounds in that:

1. It possesses a validity test by which an hypothesized model may be rejected, if false. That is, it has an error theory.

2. It can handle continuous, interval-level numerical values as well as ordinal data.
3. It provides an efficient way of diagnosing and testing both marginal and joint relationships without resort to scaling things one-at-a-time.
4. The statistical methodology is well-understood, easily applied and robust. It is the same methodology for testing significance and estimating parameters for general linear statistical models.

Functional Measurement Methodology

We will illustrate the manner in which the methodology is applied with reference to adding and multiplying models (1 and 4). Suppose we hypothesize that consumer trips by bus are conditional upon fare (F), frequency of service (S) and distance to closest stop (D). We can then write:

$$E_{ij} = k + w_1 F_i + w_2 S_i + w_3 D_i + E_{ij}, \text{ or} \quad (5)$$

$$E_{ij} = k' + k'' \pi F_i S_i D_i + E_{ij} \quad (6)$$

where E_{ij} is the overall evaluation given by consumer i to alternative j , the k 's are scaling constants, w 's are weight parameters and E_{ij} is an error term assumed to be normally distributed with zero mean and unit variance.

Equation 5 is testable by having each consumer (i) make a numerical evaluation of each alternative (j). How this is done will be ignored because extended treatments are available elsewhere (Anderson, 1974, 1976; Louviere, 1978). We assume that E_i is some response or judgment (decision, degree of preference, etc.) observed on an interval scale. The only restriction, which may be considerably relaxed, is that the j alternatives be chosen such that they constitute the cells of a factorial or fractional factorial combination of the experimental variables. Again, standard treatments of experimental design are available and will not be elaborated here (see, e.g., Winer, 1962; Snedecor and Cochran, 1968). We will agree to term a complete design with all possible combinations of experimental factors a full factorial (FF); while designs that are fractionated and have some effects aliased or confounded with other effects we will term fractional factorial designs (FFD).

If the j alternatives are selected by means of an FF or FFD plan, then equations 5 and 6 may be tested. Either analysis of variance or multiple linear regression may be used to test these equations. These equations predict that certain effects (parameters) will be significant or nonsignificant. In particular, equation 5 predicts that only the main effects will be significant, while equation 6 predicts that both main and interaction effects will be significant. The pattern of these effects is critical because alternative algebraic formulations imply alternative patterns (see Anderson, 1974, 1976). Detailed treatments of the testing and/or diagnosis of these models are available elsewhere (see Anderson, 1974, 1976 for example) and will not be pursued here.

All of these tests are complemented by graphical plots of the data: Briefly, equation 5 predicts that all main effects plots (graphing the marginal means against their experimental values) will be linear and that all interaction plots will consist of a series of parallel straight lines, with slopes equal to the corresponding w 's (see Anderson, 1974, 1976 for detailed discussion). Similarly, equation 6 predicts that the main effects plots will be the same as equation 5, but the interaction plots will not: they will consist of a series of diverging straight lines (see Anderson, 1974, 1976 for a detailed treatment). Both the statistical and the graphical evidence must agree. Ordinarily both are

employed in diagnosis and testing. We will now illustrate the application of the approach in a series of brief empirical examples. These examples include: 1) consumer evaluations of alternative bus systems to satisfy general transportation needs; 2) consumer evaluations of alternative carpooling scenarios; 3) consumer evaluation of alternative grocery shopping destinations; and 4) consumer evaluation of alternative towns as places to live.

Empirical Applications

Bus System Evaluations. Beginning in 1972, a series of studies was undertaken by Louviere, Norman, Levin and their associates to examine the process by which consumers trade-off attributes of public bus systems. The results are summarized elsewhere (Norman and Louviere, 1974; Louviere and Norman, 1977; Norman, 1977, Levin and Gray, 1977; and Levin, 1978) and will be briefly restated here.

In an initial study of consumer trade-offs, using students as subjects at the University of Iowa, it was found that the students both individually, and as a group, acted "as if" they employed a multiplicative process in evaluating "how frequently" they might use 27 alternative bus systems that differed in fare, headways and walking distance to closest stop. It was also demonstrated that the marginal relationships between these factors and the subjective response was nonlinear, describable by power functions of the form $y = a + bx^c$.

This study was followed by a second in which trip purpose, fare, walking distance, hours of operation and number of stops between origin and destination were covaried in a series of sub-factorial designs according to the method of differential information developed by Norman (1976). For each trip purpose, subjects evaluated four three-factor subdesigns. This permits uniquely separable estimates of the weight parameters in the context of a geometric averaging model, which is a multiplicative model. Results supported the model and revealed that the weights on each factor differed by trip purpose: Fare had a higher weight for shopping than work trips; walking distance had a higher weight for work trips than shopping trips; the number of intervening stops had a greater effect than hours of operation for work trips, but the reverse was true for shopping trips. Additional studies (e.g., Wilf, 1974 and Louviere, Meyer, Stetzer and Beavers, 1974) have also supported multiplicative evaluation processes. Unfortunately, all of these preliminary studies were conducted with college students.

Two recent studies of nonstudent employees at the University of Iowa, however, have yielded somewhat different results (Meyer, Levin and Louviere, 1978). These studies manipulated time difference of bus over car, cost difference of car over bus and number of other passengers. Thus, subjects were given situations to evaluate described by different excess travel time by bus, higher costs by car, and number of other individuals present (to simulate drive alone, car pool, or transit).

A cluster analysis was used to define homogeneous groups of subjects according to the pattern of their responses to the alternative situations. The clusters grouped respondents who favored the car, were neutral and who favored the bus (The response scale used was a 150 millimeter line, marked off by "certain to take car" or "certain to take bus" at either end. Clusters included subjects who responded toward the car end, the bus end, or in the middle, on the average). Graphs of the data from each group revealed that only car biased subjects exhibited multiplicative evaluation processes. Both other groups exhibited processes best described by the

differentially weighted averaging model of equation 3. However, more research is required to extend the generality of these findings past the experiment, its context and this group of subjects.

In another phase of these Iowa studies, the authors fit a modified regression equation to actual bus patronage behavior of the employees. Respondents were asked to estimate the number of bus trips made to work in the previous month, the distance from their residence to work, whether bus or car was more expensive for the trip and whether they felt that the bus was really available to them to make the trip (yes/no). From the experimental data, a measure of bias toward one or the other mode can be derived: it is each respondent's grand mean response over all evaluations. It was hypothesized that this measure captures a wide array of attitudinal effects and that it would have a significant effect on actual (reported) patronage. The following model was tested:

$$\begin{aligned} \text{Proportion of trips by bus} = & (b_1 \text{ Grand Mean} + b_2 \text{ Home--} \\ & \text{Work Distance} + b_3 \text{ Is Car} \\ & \text{Cheaper (1 = Yes)}) \times \\ & \text{Whether Bus Available} \\ & (\text{No} = 0) + E \end{aligned} \quad (7)$$

This equation accounts for almost 80% of the variation in the data from 97 respondents, with the grand mean factor (bias) alone accounting for almost 70%. A second analysis in which respondents were classed by usual mode and employing equation 7 in a binary discriminant analysis permitted 94% correct classification, a figure that compares well with more traditional disaggregate models such as the multinomial logit (Domencich and McFadden, 1975). A missing element in these studies has been an analysis of the extent to which group membership can be predicted by sociodemographic and environmental measures for each individual. The results strongly suggest that if individual grand mean responses are predictable from forecastable socioenvironmental and demographic variables, mode choice can be readily forecast. Work is proceeding to test these notions on several projects, but is not available at this writing.

Consumer Carpooling Decisions. Traditional analyses of factors in carpooling have avoided nonquantitative factors such as interpersonal variables because of difficulties in their measurement. However, recent studies by Levin, Mosell, Lamka, Savage, and Gray (1977) and Mosell, Lamka, and Levin (1977) have included factors such as the sex of potential carpool riders and whether or not the respondent had a prior acquaintance with other potential riders. The functional measurement approach allows an analysis of how each factor combines with each other factor to determine the potential desirability of a carpool.

Results indicate that interpersonal factors have an effect on judged carpool desirability comparable to the effects of more traditional time and cost factors. In particular, the desirability of an alternative carpool description can be described as an average of the desirability of the individual potential riders, and the desirability of an individual potential rider can be described as a multiplicative function of sex and acquaintanceship. Thus, if the potential rider is an acquaintance, the sex of the rider is of little or no consequence. If the rider is not an acquaintance, sex is important, and both male and female respondents prefer a female rider. Acquaintanceship by itself appears to have a dramatic effect on carpool desirability: Extremely low ratings are given to carpools with no acquaintances, but if even one potential rider is identified as an acquaintance, the desirability of carpooling increases dramatically. This finding is currently

being pursued as a means for incorporating interpersonal factors in carpool promotional programs with the aid of funding from the Urban Mass Transit Administration.

Consumer Evaluation of Alternative Grocery Shopping Destinations. Studies of samples of the general population of Tallahassee, Florida and Laramie, Wyoming have applied the Functional Measurement approach to the identification of consumer evaluation functions and the use of these functions to predict actual store patronage. On the basis of preliminary survey work in both cities, it was found that consumers apparently traded-off perceived prices, selections, and convenience in choosing real supermarkets. A Functional Measurement experiment which varied levels of these three factors in a 3³ FF plan was evaluated by random samples of citizens in the two cities drawn from telephone directories.

Data analysis confirmed the a priori hypothesis that consumers would act "as if" they employed a multiplicative evaluation process in judging stores described by combinations of the three factors. Respondents were asked to supply a numerical evaluation of "how good" each of a list of real supermarkets with which they felt familiar was on each experimental factor (prices, selection, convenience). The judgment scale was the same as that employed in the experiment--a 20 category rating scale. The average category value assigned to each real supermarket on each factor is its expected scale value on that factor if the experimental results are approximately true. Then these values may be combined according to the following rule:

$$E_i = p_i \cdot s_i \cdot c_i \quad (8)$$

where E_i is the expected overall evaluation or "utility" associated with each of the i ($= 13$ in Tallahassee; $= 6$ in Laramie) real supermarkets and p_i , s_i and c_i are, respectively, the average value of price, selection and convenience ratings. Then, by measuring either patronage (Tallahassee) or gross yearly sales (Laramie) we hypothesize:

$$B_i = f(E_i) \quad (9)$$

where B_i is the measure of patronage or sales and E_i is defined in 8. In both cities B_i was linearly related to E_i : The correlation was .93 in Tallahassee and .96 in Laramie. Again, more research is needed to extend the generality of these findings.

Consumer Evaluation of Alternative Towns as Places to Live. Resource exploitation is proceeding rapidly in the Rocky Mountain West and a major concern is understanding and predicting which towns nonlocal employees will choose for their residences when they are brought in to western industrial sites. One simple hypothesis is that they will trade-off the amenities offered by the town with commuting distance to work. It is now well-known that the number of various types of goods and services offered in central places is linearly related to the population base of those places (see, e.g., Berry, 1967). Hence, a range of 10 services were predicted from population sizes of 250, 500, 1000, 1500, 2000 and 2500 and were arranged in a series of 36 different paragraph descriptions which listed the population, the number of each type of good and service available and one of six (15, 30, 45, 60, 75, 90 miles) different commuting distances. This is a 6 x 6 FF plan because amenities are perfectly linearly related to population and cannot be statistically separated.

Seventy-five staff, faculty and student volunteers at the University of Wyoming judged "how desirable" each town alternative would be to them as a place to live if they had a job at an industrial site x miles away in Western Wyoming. Analytical results confirmed the

a priori hypothesis that subjects would act "as if" they processed the information in a multiplicative manner, similar to a gravity model. The following equation was found to describe the data well:

$$E_i = ae^{-b\text{Distance}_i} (c + d\text{Population}_i^p) \quad (10)$$

where a, b, c, d and p are empirical parameters and E_i is the observed average evaluation given to the i towns ($i = 36$).

A set of real choice proportions for seven different industrial sites in three states was available to test the model (Old West Commission, 1975). Let p_i be the observed proportion of nonlocal employees from a particular plant that chose town i as a residence. Because the commuting distances and populations of the i towns can be approximated, they may be substituted in equation 11 to yield \hat{E}_i --the expected evaluation or utility which the experimental subjects would have assigned the i towns had those distance and population combinations been used in the experiment. We then hypothesized that $p_i = f(E_i)$.

The graph of p_i vs. E_i is highly monotonic and can be approximated by a function of the form, $p_i = a'e^{b'E_i}$. The correlation between the two sets of observations is .92 ($i = 56$). It is clear (at least in this study) that a simple transformation exists to map responses or utilities into expected choice proportions. A model may then be re-estimated on expected choice proportions and used to forecast the number (or proportion) of employees at plant i that will choose town j as a residence.

Summary

This paper has reviewed and briefly described and discussed theory and methodology from the Functional Measurement approach to modeling consumer information processing in judgment. It was demonstrated that consumer trade-off functions can be readily derived for a wide class of applied problems and that these functions are related to actual consumer behavior and can be used to forecast same. A considerable amount of research remains, however, to develop an understanding of the generality of the models across samples, places and individual consumer groups. Moreover, little is known about the effects, if any, of social, demographic and environmental factors on the individual decision function. Such knowledge is necessary to implement the models for practical forecasting problems.

Perhaps the most important conclusion is that individual decision processes can be diagnosed and modeled using simple experimental designs and statistical tools. Models are derived empirically and many applications have shown the typical a priori assumption of linear forms to be inadequate to describe real-world behavior. This extends work that began in the laboratory in 1962 to real field problems and demonstrates that the approach can be most useful in understanding consumer travel decision processes.

References

- N. H. Anderson, "Functional Measurement and Psychophysical Judgment," Psychological Review, 77(1970), 153-170.
- N. H. Anderson, "Social Perception and Cognition," Technical Report CHIP 62, Center for Human Information Processing, University of California at San Diego, La Jolla, California, June, 1976(a).
- N. H. Anderson, "How Functional Measurement Can Yield Validated Interval Scales of Mental Quantities," Journal of Applied Psychology, 6(1976), 677-693.
- N. H. Anderson, "Note on Functional Measurement and Data Analysis," Perception & Psychophysics, 21(1977), 201-215.
- B. J. L. Berry, Geography of Market Centers and Retail Distribution, Englewood Cliffs, N. J.: Prentice-Hall, 1967.
- T. Domencich and D. McFadden, Urban Travel Demand: A Behavioral Analysis, Amsterdam: Elsevier (North Holland), 1975.
- I. P. Levin, "The Development of Attitudinal Modelling Approaches in Transportation Research," in D. A. Hensher & P. R. Stopher (Eds.), Proceedings of the Third International Conference on Behavioural Travel Modelling, Pergamon Press, (1978) forthcoming.
- J. J. Louviere, "Psychological Measurement of Travel Attributes," in D. A. Hensher and M. Q. Dalvi (Eds.), Determinants of Travel Choice, London: Teakfield, Farnborough (Saxon House Studies), 1978, in press.
- J. J. Louviere, R. J. Meyer, F. Stetzer, and L. L. Beavers, "An Experiment to Derive Predictive Models of Public Response to Policy Manipulations in Public Bus Transportation," Technical Report No. 35, Institute of Urban and Regional Research, The University of Iowa, Iowa City, January, 1974.
- J. J. Louviere and K. L. Norman, "Applications of Information Processing Theory to the Analysis of Urban Travel Demand," Environment and Behavior, 9(1977), 91-106.
- R. J. Meyer, I. P. Levin, and J. J. Louviere, "Functional Analysis of Mode Choice," Paper submitted for presentation at the 1978 meetings of the Transportation Research Board, Washington, D.C., January, 1978.
- M. K. Mosell, C. M. Lamka, and I. P. Levin, "How Do I Get There from Here? Attitudes Toward Different Modes of Transportation," Proceedings of the Iowa Academy of Science, (1977) forthcoming.
- K. L. Norman, "A Solution for Weights and Scale Values in Functional Measurement," Psychological Review, 83(1976), 80-84.
- K. L. Norman, "Attributes in Bus Transportation: Importance Depends on Trip Purpose," Journal of Applied Psychology, 62(1977), 164-170.
- K. L. Norman and J. J. Louviere, "Integration of Attributes in Public Bus Transportation: Two Modeling Approaches," Journal of Applied Psychology, 59(1974), 753-758.
- Old West Commission, Construction Worker Profile, Billings, Montana, 1975.
- J. Shanteau, and C. M. Troutman, "Commentary on Bettman, Capon, and Lutz," Journal of Consumer Research, 1(1975), 16-18.
- P. Slovic, B. Fischhoff, and S. Lichtenstein, "Behavioral Decision Theory," Annual Review of Psychology, 28(1977), 1-39.
- P. Slovic, and S. Lichtenstein, "Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment," Organizational Behavior and Human Performance, 6(1971), 649-744.

G. W. Snedecor and W. G. Cochran, Statistical Methods, Ames, Iowa: Iowa State University Press, 1968.

C. M. Troutman and J. Shanteau, "Do Consumers Evaluate Products by Adding or Averaging Attribute Information?" Journal of Consumer Research, 3(1976), 101-106.

C. M. Wilf, Predicting Potential Bus Patronage: An Integration Theoretic Approach, Unpublished M.A. Thesis, Florida State University, Department of Geography, June, 1974.

W. L. Wilkie, and E. A. Pessemir, "Issues in Marketing's Use of Multi-Attribute Attitude Models," Journal of Marketing Research, 10(1975), 428-491.

B. J. Winer, Statistical Principles in Experimental Design, New York: McGraw-Hill, 1962.

EXTENDED CONJOINT ANALYSIS WITH INTENSITY MEASURES
AND COMPUTER ASSISTED INTERVIEWS:
APPLICATIONS TO TELECOMMUNICATION AND TRAVEL

John R. Hauser, Northwestern University
Steven M. Shugan, University of Rochester

Abstract

This paper reports two advances in the measurement of consumer preferences. First, a general theory extends conjoint measurement from ordinal measures (rank order only) to more effective and efficient intensity measures (interval, ratio, or hybrid). Tests are given to identify the appropriate theory to describe consumer response and procedures are given to estimate the preference functions. Second, a marketing information system (P.A.R.I.S.) is described which can automatically encode any questionnaire on an interactive computer system for computer assisted interviewing. P.A.R.I.S. has special commands for the preference measurement questions; it automatically stores all data in special files for easy access and has a subroutine for the improved conjoint analysis. The general theory as implemented through P.A.R.I.S. is illustrated with a case application to the analysis of consumer preference for new telecommunications technology. Comparisons are made to standard conjoint analysis and to preference regression.

Introduction

To design new transportation or communication services, managers and analysts must understand how consumers will react to such services and must be able to predict the subsequent usage of those services. An important component in this analysis is the measurement of consumer preferences, i.e., the measurement of how consumers value the various attributes of transportation or communications service and how they aggregate these valuations to form an overall evaluation of each alternative.

A number of techniques have been used in both marketing and in transportation to measure or estimate consumer preferences. The first set of techniques, which we will call "group-level" techniques, are used extensively in transportation demand analysis and include the "dis-aggregate behavioral demand models." These techniques normally represent the preference value p_{ij} , that individual i places on alternative j as a weighted sum of i 's perception of j with respect to a set of attributes. If w_k is the "group" weight for attribute k , and x_{ijk} is i 's perception of j with respect to k , then the model is represented by $p_{ij} = \sum_k w_k x_{ijk}$. Among the techniques used are logit analysis (McFadden, 1970), preference regression (Urban, 1975), and maximum score. The advantage of these models is that they are relatively easy to estimate because they are based on statistical analysis of consumers' preference or choice among existing services. Their disadvantages are that they merge individual difference and estimate group weights (w_k) rather than individual weights (w_{ik}) and that the simple linear form may not accurately represent consumer behavior (Farquhar, 1977; Hauser and Urban, 1977b, Keeney and Raiffa, 1976). Their primary use is as a preliminary screening process to identify important attributes and provide "first order" predictions.

As the design of the product or service is refined, analysts need stronger measures and an ability to better understand the mapping of perceptions into preference. Furthermore, for various segmentation strategies it is important to identify individual differ-

ences by measuring preference functions for each individual consumer. These "individual-level" techniques normally represent preference, p_{ij} , as a sum or product of non-linear functions of the attributes, i.e., $p_{ij} = \sum_k u_{ijk}(x_{ijk})$. Among these techniques are conjoint analysis (Tversky, 1967; Green and Wind, 1973), trade-off analysis (Johnson, 1974), direct utility assessment (Hauser and Urban, 1977b). The advantages of these models are that they provide more detailed identification of individual differences and that their non-linear forms are more sensitive to decreasing returns and risk aversion effects (Raiffa, 1970). Their disadvantages are that they are more expensive because they normally require a personal interview, that they must often expand the stimuli set to pseudo-products (products represented by their attribute levels) to get sufficient degrees of freedom for estimation, and that the consumer task is often tedious. When compared to group-level techniques the "individual-level" techniques are more accurate but more expensive. Because of their advantages, these techniques have definite use in the design of transportation services if used judiciously (Green and Wind, 1975; Market Facts, 1976), but there is a definite need for improvement to make them more efficient and cost effective.

This paper covers two parallel improvements in the measurement of individual-level preference functions. The first, evaluation theory, generalizes conjoint analysis to stronger, more efficient measures of preference. These intensity measures are shown in an empirical example to improve prediction relative to both conjoint analysis and a representative "group-level" technique--preference regression. The second improvement, interactive interviewing, provides for more cost-effective measurement and faster, less expensive analysis. The general theory and the interactive interviewing system are covered in detail in two separate papers available upon request from the authors (Hauser and Shugan, 1977; Shugan and Hauser, 1977).

Theory Based on Intensity Measures

Standard conjoint analysis is an effective technique to measure consumer preference, but the consumer task is quite tedious requiring the consumer to rank order 20-40 "products" (actual or pseudo-products) in terms of preference. As a result, many researchers have modified conjoint analysis to reduce the consumer task. For example, Green and Wind (1975) use a fractional factorial design to reduce the number of stimuli. Johnson (1974) uses tradeoff matrices that require consumers to rank order "products" where only two attributes vary at a time. Hauser and Urban (1977b) use von Neumann-Morgenstern theory to formulate indifference questions¹ that require consumers to set an attribute level of one product such that it is equal in preference to another fully-specified product. All three procedures simplify the consumer task, but applications still require 20-40 minute interviews.

¹Indifference questions are a limiting case of rank order questions.

Furthermore, each theory measures ordinal preference, i.e., a ranking over "products,"² rather than intensity of preference. Although a set of conjoint models over a consumer population estimates how many people choose each product, conjoint analysis does not estimate ratio, interval or probabilistic preferences. Such intensity measures are potentially better indicators of the consumer evaluation process and more accurate predictors of behavior.

If conjoint analysis could be extended to intensity measures such as dollar metric (Pessemier, 1977) or constant sum paired comparisons (Torgenson, 1958), it is reasonable to posit that more information could be gathered per question and as a result, fewer questions need be asked. (See figure 1 for an example of a constant sum paired comparison [CSPC] question.) Furthermore, if the consumer gives consistent answers, it may be possible to measure preference functions that incorporate intensity of preference. Support for this con-

FIGURE 1
An Example of Constant Sum Paired Comparison Measurement
(Respondent's answers are in italics.)

DIVIDE 100 CHIPS BETWEEN EACH OF THE FOLLOWING PAIRS OF HYPOTHETICAL DEODORANTS:

PRODUCT A	PRODUCT B
PUMP SPRAY	AEROSOL
HERBAL SCENT	UNSCENTED
REGULAR	ANTI-PERSPIRANT

ENTER CHIPS ...

74, 26

jecture comes from recent simulation and empirical results (Green, 1976; Carmone, Green and Jain, 1976; Cattin and Wittink, 1976; Hauser and Urban, 1977a) which show that conjoint analysis is robust with respect to the metric/non-metric assumption. In fact, there are definite indications that the metric models which treat ordinal data as interval data can outperform their non-metric counterparts. Further support comes from Silk and Urban (1976) who report tremendous success with CSPC for actual products chosen in simulated purchase environment (over 80% of the uncertainty in behavior explained). The natural extension would then be to use intensity measures directly rather than arbitrarily forming interval data from ordinal data.

But conjoint theory (Tversky, 1967) as well as utility theory (von Neumann-Morgenstern, 1947) is based on ordinal preference. Before intensity measures can be used, assumptions and properties of the measures and preference functions must be identified. Tests must be developed to determine whether the consumer reacts consistently to the task and to determine what theory best explains his or her answers. This will determine whether the preference function should exhibit ordinal, interval, ratio, probabilistic, or hybrid properties.

Hauser and Shugan (1977) show that consumer reaction to intensity measures can be represented by a general equation based on property operators and measurement relations. They show that two axioms--property asymmetry and property transitivity--are necessary for consistency of the representations and a general independence property--evaluative independence--provides simple decompositions for ease of estimation. Rather than cover

²Von Neumann-Morgenstern preference functions can handle lotteries, i.e., products with uncertain or risky attributes.

the detailed derivations in this paper, we will simply state the results for four special cases: ordinal, interval, ratio, and hybrid.

Conjoint Analysis (Ordinal Theory)

Ordinal theory is based on the standard assumption that the intensity measures give only rank order information. I.e., consider the CSPC question in figure 1. Let x_k be the value of the k^{th} attribute, e.g., scent $(k=2)^j$, for the j^{th} product, e.g., product A ($j=1$). Let $x_j = (x_{j1}, x_{j2}, \dots, x_{jk})$. Let a_{ij} be the number of chips allocated to product i when comparing i and j , and let $a_{ji} = 100 - a_{ij}$. Let $u(x_j)$ be the preference function that a_{ji} maps the attributes x_j into a measure of preference, i.e., $p_j = u(x_j)$ where p_j is a scalar measure of "goodness" for product j . In this notation ordinal theory is simply stated:

$$a_{ij} \geq a_{ji} \text{ implies } u(x_i) \geq u(x_j) \quad (1)$$

Property transitivity is ordinary transitivity. I.e., if δ_{ij} is defined such that $\delta_{ij} = 1$ if $a_{ij} > a_{ji}$, $\delta_{ij} = 0$ if $a_{ij} = a_{ji}$, and $\delta_{ij} = -1$ if $a_{ij} < a_{ji}$ then property transitivity among three products-- i , j and k --is given by:

$$\delta_{ij} + \delta_{jk} - \delta_{ik} = \delta_{ij} \delta_{jk} \delta_{ik} \quad (2)$$

Under evaluative independence (which is known in ordinal theory as pairwise preferential independence [Ting, 1971; Farquhar, 1977; Keeney and Raiffa, 1976]) it can be shown that the appropriate decomposition is:

$$u(x_j) = u_1(x_{j1}) + u_2(x_{j2}) + \dots + u_k(x_{jk}) \quad (3)$$

Equations 1, 2 and 3 can also be extended to handle risky alternatives, i.e., products with uncertain attribute levels by the use of von Neumann-Morgenstern theory. Equations 1 and 2 are extended to lotteries (Keeney and Raiffa, 1976; Hauser and Urban, 1977a; Hauser, 1976) and equation 3 becomes Keeney's quasi-additive form (Keeney, 1972).

Estimation is by standard conjoint estimation (Green and Wind, 1975; Johnson, 1974; Srinivasan and Shocker, 1973) or, for risky alternatives, with direct utility assessment (Hauser and Urban, 1977b).

Intensity Measures (Interval Theory)

Conjoint analysis has proven successful in the past but ordinal theory uses only part of the information in CSPC or equivalent measures. An alternative theory, based on fundamental axioms by Shapley (1975) is interval theory. This is given simply by:

$$u(x_i) - u(x_j) = a_{ij} - a_{ji} \quad (4)$$

The test of consistency is additive transitivity.

$$(a_{ij} - a_{ji}) + (a_{jk} - a_{kj}) = (a_{ik} - a_{ki}) \quad (5)$$

Under evaluative independence, the decomposition is again additive:

$$u(x_j) = u_1(x_{j1}) + u_2(x_{j2}) + \dots + u_k(x_{jk}) \quad (6)$$

If we discretize each $u_k(x_{jk})$, i.e., define $u_k(x_{jk}) = \sum_{\ell} \lambda_{\ell k} \delta_{\ell k j}$ where $\delta_{\ell k j} = 1$ if product j has attribute k at level ℓ and $\delta_{\ell k j} = 0$ otherwise, then the estimation equation is:

$$a_{ij} - a_{ji} = \sum_k \sum_{\ell} \lambda_{\ell k} (\delta_{\ell k i} - \delta_{\ell k j}) + \text{error} \quad (7)$$

The part-worths, λ_{jk} , are estimated with ordinary least squares regress (OLS) or linear programming (Srinivasan and Shocker, 1973).

Intensity Measures (Ratio Theory)

Originally CSPC is based on a ratio assumption (Torgenson, 1958). Hauser and Shugan (1977) show that the ratio assumption can be extended from comparisons over physical products to comparisons over attribute bundles. This theory is given by:

$$u(x_i)/u(x_j) = a_{ij}/a_{ji} \quad (8)$$

The test of consistency is multiplicative transitivity:

$$(a_{ij}/a_{ji}) \cdot (a_{jk}/a_{kj}) = (a_{ik}/a_{ki}) \quad (9)$$

It can be shown that the appropriate decomposition is multiplicative:

$$u(x_j) = u_1(x_{j1}) \cdot u_2(x_{j2}) \cdot u_3(x_{j3}) \cdot \dots \cdot u_K(x_{jK}) \quad (10)$$

If we discretize $u_k(x_{jk})$ as an interval theory, then under the logarithmic transformation the estimation equation becomes:³

$$\log(a_{ij}/a_{ji}) = \sum_k \sum_l (\delta_{lki} - \delta_{lkj}) \log \lambda_{lk} + \text{error} \quad (11)$$

The part worths, λ_{lk} , are obtained from estimates of $\log \lambda_{lk}$ based on OLS or linear programming.

Hybrid Theories

Given the general form it is possible to extend the above theories. One that we have found most useful is a hybrid theory to investigate non-linearities at the extreme end of the scales. The form is a combination of the ratio and interval theories:

$$u(x_i) - (a_{ij}/a_{ji})^\gamma u(x_j) = \beta(a_{ij} - a_{ji}) \quad (12)$$

Note that if $\gamma \rightarrow 0$, $\beta \rightarrow 1$ the interval theory applies, if $\gamma \rightarrow 1$, $\beta \rightarrow 0$ the ratio theory applies. In between estimation is based on non-linear estimation routines (Abelman, 1976; Cohen, 1975). Note also that if $\gamma = 0$ or $\beta = 0$, linear techniques apply.

Tests to Identify Appropriate Theory

The above four theories are representative of the possible theories that can result from the general form.⁴ They illustrate the breadth of possible procedures to estimate preference functions based on CSPC questions. To proceed further we must devise a test to empirically distinguish how each consumer reacts to the CSPC questions. With this test we can first observe how the consumer responds to a given battery of questions and then based on his or her responses branch to the appropriate theory and estimation. To perform this test, consider all product triplets and define a_{ik}^I as the closest integer such that the interval test, equation 5, holds. Define a_{ik}^R as the closest integer such that the ratio test, equation 9, holds. These differences minimize the scale differences inherent when comparing "errors" in equations 7 and 11. Suppose that a_{ik} is the

³ The appropriate functional representation is multiplicative in the λ_{lk} terms raised to the δ_{lkj} power.

⁴ Hauser and Shugan (1977) cover other theories such as stochastic preference which includes an individual level logit model.

respondent's actual answer, then define the following absolute error tests:

$$T_I = (1/n) \sum_t | a_{ik}^I - a_{ik} | \quad (13)$$

$$T_R = (1/n) \sum_t | a_{ik}^R - a_{ik} | \quad (14)$$

where the summation is over all possible tests, t , within a design and n is the number of such tests. If $T_I \leq T_R$, then the consumer is more likely responding via the interval theory. If $T_I < T_I$, then it is more likely the ratio theory applies. If both T_I and T_R are above some cutoff, T_c , reject both theories. (In practice we have found $T_c = 20$ is a good cutoff. Furthermore, we have found that the root mean square error modification of equations 13 and 14 give similar results.)

Summary

This section has reviewed the results of a theory based on intensity of preference. In particular it has stated the appropriate functional forms and has given equations to estimate preference functions via each theory. Finally, it has given a test to distinguish from a set of data, which theory best describes how consumers are reacting to the CSPC scales.

As stated earlier, these theories have great potential for accuracy because they utilize more information in consumer judgments than do the existing methods of conjoint analysis and preference regression. This remains to be shown. Following a discussion of computer assisted measurement, the section on application will give empirical evidence to support this conjecture, but first we will digress to illustrate how one can practically measure such functions.

Computer Assisted Consumer Measurement and Data Analysis

The general intensity theory makes possible the estimation of preference functions based on CSPC data. Because we expect that more information is contained in this data, we expect that the general theory will lead to improved accuracy or conversely, the same accuracy but with fewer questions. But this theory depends on our ability to quickly and efficiently ask CSPC questions. To do this we developed a marketing information system, P.A.R.I.S., which can implement a general questionnaire but which has special capabilities built in to implement CSPC questions. Furthermore, P.A.R.I.S. automatically encodes the data and sets up computer files for easy access and analysis. Special subroutines access the stored CSPC data and estimate the preference functions. The details of the P.A.R.I.S. system are contained in Shugan and Hauser (1977). We will describe here some of the features as they relate to preference measurement.

Development of an Interactive Program

The P.A.R.I.S. system allows the researcher to develop an interactive questionnaire as easily as a written question might be developed. The researcher would accomplish this development in three steps.

First, the researcher constructs a questionnaire using the P.A.R.I.S. language. This language consists of simple commands to print questions, record answers, check ranges, branch based on consumer response, etc. The language, which consists of over thirty commands, is sufficiently general to implement most market research questionnaires.

⁵ Preference Assessment and Retrieval Information System

Second, the written questionnaire is either punched on ordinary data cards or written on-line to computer storage (tape or disc). A single command activates the P.A.R.I.S. Q-compiler which converts the human-oriented language into an alpha-numeric, machine-oriented language and sets up the appropriate files to record and decode answers and sets up a system to time and record how long it took each consumer to answer each question. The Q-compiler also checks the questionnaire for coding errors and provides a summary of each question's status, thus alerting the researcher of possible errors in questionnaire design.

Finally, the compiled version of the questionnaire is automatically input to a "mass storage" program. This program stores the questions in a format allowing efficient computer access to any question in any order. The questionnaire is now ready for implementation.

Development of a Dynamic Market Research Data Base

The actual administration of the questionnaire is accomplished either by seating a respondent at a portable terminal (e.g., Cathode-Ray tube CRT) or by allowing an interviewer to interactively record answers while obtaining them in person by telephone. In either case, P.A.R.I.S. adds the response to the data base together with the time and date when entered. The computer then provides instantaneous range-checking to insure each response is in the legitimate range allowed for that response. Illegal responses (e.g., "yes" when his/her age is asked) may be followed by a gentle computer response informing the respondent or interviewer of an error and providing a clarifying instruction. Once a legitimate response is obtained, the range of that response can determine the next question. For example, if the respondent can only evaluate three brands of deodorants, perception questions about non-relevant brands can be avoided. This branching permits very efficient questioning, minimizing the actual number of questions asked to the most relevant questions for the consumer.

When the interview is completed, summary statistics are automatically provided. Management can access all interviews to date or some selected portion of them. Statistical analysis can be performed periodically and selectively. A master file allows a researcher to determine at a glance the current sample size, progress for the entire study, and how long each questionnaire administration took. More detailed information, e.g., how long a partial question's administration took and the answer given, can be obtained from the main database record. A special comment file records qualitative responses for easy access and analysis. We have found record and comment information, together with the feature that all answers are recorded regardless of whether a question is reasked because of an improper response, is very valuable for the development of test questionnaires. All answers, including mistakes, can be accessed to fully pretest a questionnaire.

Special Constant Sum Paired Comparison Commands

P.A.R.I.S. is designed with special commands for preference analysis by conjoint or evaluation theory. For example, a single command "READ n CHIPS" sets up a constant sum paired comparison question with automatic range checks and tests to ensure that the consumers' responses sum to n. If the responses do not sum to n, or if negative or non-numeric answers are given, the system diagnoses the problem and informs the respondent of his or her mistake. The form for quick answers or a longer form with more explanation. See figure 1 for an example of the short form. Furthermore, the system automatically sets up a computer routine to encode the constant sum response and place it in the record file

for easy access. This encoded data is input to a related subroutine which uses linear programming to estimate the preference functions. This subroutine is a modification for constant sum data of ideas expressed by Srinivasan and Shocker (1973) in LINMAP. Experience to date has shown that the constant sum paired comparison task for preference measurement is readily accepted by consumers.

Typical Questionnaire

The next section reports an empirical application of a typical questionnaire and gives an example (figure 2) of what the consumer sees. We will put off this discussion until that section.

Empirical Example: Telecommunications Innovations⁶

Previous sections presented the results of a general theory for intensity of preference and described a marketing information system to implement that theory. We present here an empirical example of the use of P.A.R.I.S. to estimate intensity of preference functions. The remainder of this section is taken from section 3 of Hauser and Shugan (1977). The entire paper is available from the authors.

The empirical problem is to design a mix of telecommunications technology for use in a small research community. Scientific research requires effective communication among scientists, but in many government laboratories cooperating scientists and managers find themselves in buildings 2-3 miles apart. Furthermore, laboratories, such as Los Alamos Scientific Laboratory (LASL) in New Mexico, do much of their work for federal agencies and there is a strong need for effective communication with managers and policy makers in Washington, D.C. Currently, the most common means of communication are telephone (39% in LASL) and personal visit (58% in LASL), with only a small percentage (3% in LASL) of the interactions using other means. The National Science Foundation would like to enhance communication among the scientists, managers, and policy makers with an improved system that is more effective than telephone for technical communication, yet more efficient than personal visit in terms of cost, time, and energy. Among the options being considered are closed circuit television, telecopiers (facsimile transfer devices or teletypewriters), and narrow-band televideo systems (an attachment to the telephone which transmits still pictures over voice-grade telephone lines). But since the laboratories have limited budgets, each laboratory would like to implement the communications system that would be most cost-effective. To do this the laboratories need to know how scientists and managers would react to the various systems.

Study Design. To address this problem we used the normative methodology described in Hauser and Urban (1977a) to identify the relevant dimensions that describe communications options and to identify the relative importances of these dimensions. These dimensions form the basis for the CSPC questions used in the estimation of preference functions.

First, consumer focus groups were run and analyzed to produce an indication of the choice process, consumer semantics, and a set of 25 attribute scales to character-

⁶ The majority of this section is taken from John R. Hauser and Steven M. Shugan, "Efficient Measurement of Consumer Preference Functions: A General Theory for Intensity of Preference," Working paper, Department of Marketing, Northwestern University, Evanston, Illinois, July, 1977.

ize consumer reactions to communication technology. Based on the focus groups and on previous research in the area of communications, a mail-back questionnaire was designed and implemented in which consumers rated telephone, personal visit, and the three new communications options (1-page concept statements) on the 25 attribute scales. Factor analysis of the response revealed two perceptual dimensions labeled "ease of use" and "effectiveness." Ease of use correlates with the ability to find the right person, save time, eliminate paperwork, and get a quick response as well as saving hassle, planning, time and cost; effectiveness correlates with the ability to exchange scientific and technical information, persuade people, convey all forms of information, control the impression you want to make, monitor people, operations, and equipment, yield a high level of human interaction, solve problems, express feelings, and enhance idea development.

Scientific and managerial communication is complex and it is probable that the communications needs would vary by individual depending upon his or her requirements. Thus, to accurately analyze preferences for communications options, we need to stratify by use scenario (purpose, distance between communicators, relation of communicators, etc.) and estimate preference functions within each category. Evaluation theory is used to measure these preference functions.

Consumer Measurement. Based on the results of the mail questionnaire, a preference assessment questionnaire was designed to measure the CSPC data needed for the preference functions. This questionnaire was then implemented via P.A.R.I.S. to scientists and managers at IASL and practicing managers enrolled in Northwestern University's Managers Program (evening work toward a master's in management). The questionnaire contained six sections: (1) warmup questions, (2) questions to establish a scenario for usage, (3) consumer rating of effectiveness and ease of use for the existing options and concept statements, (4) the CSPC questions, (5) preference ranking and usage intent for the existing products and the concepts, and (6) various personal and demographic questions and comments.

Example questions of section 2 are shown in figure 2. Review figure 1 for an example of a CSPC question. Note that the "questionnaire" handles out of range responses by gently informing the respondent of his mistake and asking for a new response.

Section 3 of the questionnaire was included to acquaint consumers with the measurement scales for effectiveness and ease of use and to provide us with their perceptions of each product or concept. Section 5 provided preference measures for the actual products and concepts. These measures were used to validate predictions made by each theory based on CSPC questions.

The complete questionnaire contained 96 questions including 16 CSPC questions and took about 15-30 minutes to complete. The administration cost including on-line hookup was about \$1 per respondent on a CDC-6400 (\$510 per cpu hour). The comparative results reported below are based on the sample of 41 practicing managers.

Results. Figure 3 gives the perceptual maps positioning of the five stimuli based on factor scores (mail survey) and the direct measures (preference survey). The close agreement of the relative stimuli position in the two maps supports the direct measures of effectiveness and ease of use as sufficient for the preference analysis. (Note that teletype terminals, used in the mail questionnaire, was revised to facsimile transfer devices in the preference questionnaire.)

FIGURE 2
An Example of Interactive Computer Measurement
by P.A.R.I.S.
(Respondent's answers are in italics.)

WE WOULD LIKE TO KNOW ABOUT YOUR MOST RECENT INTERACTION WITH A COLLEAGUE, OR A VENDOR, ETC., TO DISCUSS A PROBLEM ON WHICH ONE OR MORE OF YOU IS PRESENTLY WORKING. PLEASE CONSIDER INTERACTIONS ONLY WITH THOSE PEOPLE WHO DO NOT WORK IN THE SAME BUILDING AS YOU AND DON'T CONSIDER CALLS JUST TO SET UP APPOINTMENTS.

1. IN ADDITION TO YOURSELF, HOW MANY PEOPLE PARTICIPATED IN THE INTERACTION: (PLEASE TYPE IN THE NUMBER OF PEOPLE.)

?3

2. DID YOU USE:

- 1 = TELEPHONE
 - 2 = INTEROFFICE MEMO
 - 3 = MAIL
 - 4 = TELETYPE OR TELECOPIER
 - 5 = PERSONAL VISIT (YOU WENT TO HIM [THEM])
 - 6 = PERSONAL VISIT (HE [THEY] CAME TO YOU)
 - 7 = PERSONAL VISIT (CONFERENCE ROOM, AUDITORIUM, ETC.)
 - 8 = OTHER
- (PLEASE ANSWER WITH A NUMBER 1 THROUGH 8.)

? 10

(PLEASE ANSWER WITH A NUMBER 1 THROUGH 8.)

?8

PLEASE SPECIFY

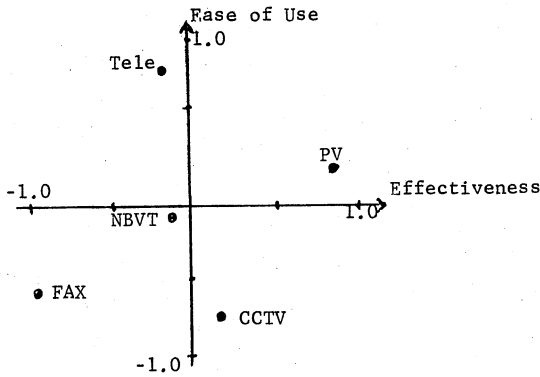
? TELEGRAM

The heuristic T-test with a cutoff of $T = 20$ indicated that 51.2% of the consumers were ratio, 31.2% interval, and 17.1% at most ordinal. The median minimum T_r , T_R was 13.1 with an interquartile range of 10.6-17.9. Based on the T-test we would expect that models based on one of the intensity of preference models would outperform standard ordinal estimation (conjoint analysis).

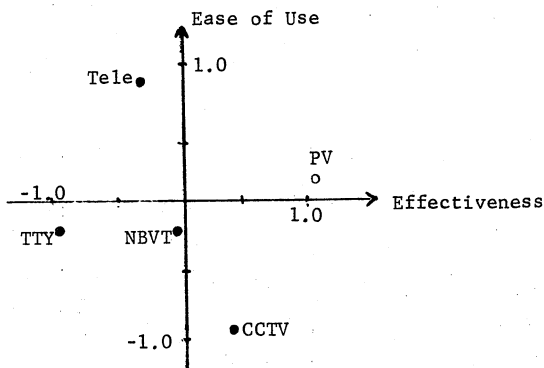
Table 1 compares first preference predictions based on linear programming estimates for the ratio, interval, and ordinal theories. For completeness, these predictions were compared to preference regression (Urban, 1975) which is an aggregate technique that assumes the same preference function for all consumers. Preference regression is widely used in the marketing literature and is representative of aggregate models such as aggregate monotonic regression (Hauser and Urban, 1977b) and logit analysis (McFadden, 1970). Recent studies have shown that predictions based on these three models are similar.

Inspection of table 1 shows that the intensity of preference theories do improve predictions over the existing models--preference regression and conjoint analysis. Furthermore, these predictions are quite good--72% of those products or concepts predicted first were indeed preferred. Detailed analysis of individual predictions reveals that the improved prediction comes primarily because the intensity of preference theories can discriminate between products that the ordinal (conjoint) theory predicted as being equal in preference. (E.g., the ratio theory might predict telephone as first preference while conjoint analysis would predict that

FIGURE 3
Perceptual Maps of Communications Options



a) Direct Measures (Standardized)



b) Factor Scores (Standardized)

Tele: Telephone NBVT: Narrow Band Video Telephone
PV : Personal CCTV: Closed Circuit Television
Visit FAX : Facsimile Transfer Device
TTY : Teletype

TABLE 1
Comparison of Models Based on Ability
to Predict First Preference

	1st	2nd	3rd	4th	5th
Conjoint Analysis (Ordinal)	57.1	25.8	12.0	4.3	0
Intensity Measure (Interval)	69.1	24.8	5.3	0.8	0
(Ratio)	71.6	22.4	5.3	0.8	0
Preference Regression	65.2	18.5	7.1	4.6	4.5

Most Preferred Product was Predicted (%)

telephone was tied with NBVT for first preference.) Thus conjoint analysis is a "correct" representation of consumer preference, but the intensity of preference theories produce preference functions that can better discriminate among products.

The comparison of preference regression and conjoint analysis is mixed. Preference regression is better at recovering first preference, but, as with any aggregate technique, it makes large errors on some individuals. Thus, table 2 compares the techniques on their ability to correctly predict market preference shares. Again the intensity of preference models provide improved predictions relative to both existing techniques. Note

TABLE 2
Comparison of Predicted Market Shares of Preference
(NBVT - Narrow Band Video Telephone, CCTV - Closed
Circuit Television, FAX = Facsimile Transfer Device)

	Tele.	Pers. Visit	NBVT	CCTV	FAX	Mean Absolute Error
Conjoint Analysis (Ordinal)	28.0	38.2	13.8	13.0	8.1	6.90
Intensity Measure (Interval)	32.5	39.0	13.0	10.5	4.9	4.54
(Ratio)	32.5	39.0	14.2	9.3	4.9	4.54
Preference Regression	31.7	31.9	12.8	16.5	7.1	7.72

(Actual) (36.6) (46.3) (9.8) (4.9) (2.4)

Predicted Market Share (%)

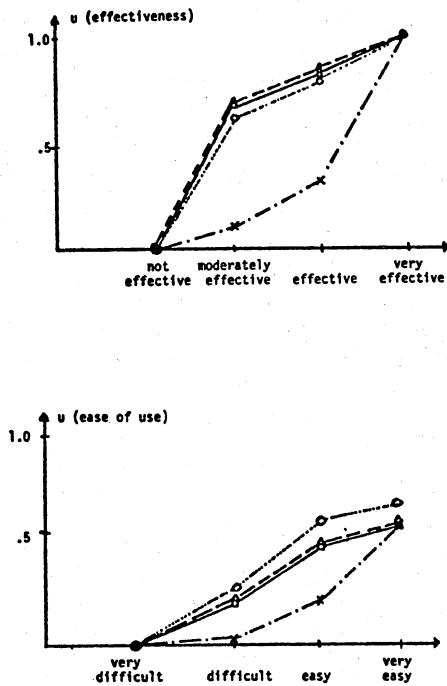
that conjoint does better than preference regression--probably because the effect of ties is less pronounced on market shares. All models do well compared to the 17.2% error that results from random prediction.

The over prediction of preference for the concepts results from the effect that consumers tend to choose an existing alternative when it is tied with a concept in predicted preference. Future research will expand evaluation theory to include this "preference inertia" effect (Neslin, 1976). Empirically when the models are applied to existing alternatives only, the mean absolute error is reduced to 1.4% for ratio. Furthermore, all relative comparisons remain the same.

Predictive accuracy is important for the evaluation of alternative products, but to improve the design of new products managers need diagnostic information to help them understand consumer preferences and thus design improved products. Although preference functions are estimated at the individual level, it is useful to present summary statistics (mean, variance, median, interquartile range) to represent the population. Figure 4 is a graph of the average preference functions. (The "utility" of effectiveness is scaled 0 to 1 for consistent comparison.) Note that the average individual functions (ratio, interval, ordinal) are quite close, although relative to the others the ordinal theory overestimates the importance of ease of use. The interpretations are quite intuitive with decreasing returns on effectiveness and a slight threshold on ease of use. These results are consistent with the above conclusion that conjoint analysis provides "correct" representations of consumer preference and that the improvements achieved by intensity of preference theories result from improved discrimination among individual preference functions.

On the other hand, preference regression gives counter-intuitive increasing returns functions, although its estimate of relative importance (ratio of maximum utilities) is consistent with the individual level preference theories. This result coupled with the predictive results suggests that the individual level measurement can better explain the details of consumer preference and identify individual differences, but preference regression, which could be done without the second survey, is adequate for a first pass at analyzing the data.

FIGURE 4
Comparisons of Preference Functions



Δ-----Δ Interval
 □-----□ Ratio
 o.....o Ordinal (Conjoint Analysis)
 x-----x Preference Regression

The final question is whether $T_o = 20$ is a good cutoff for the heuristic error measure. If the CSPC responses from consumers with $T_o > 20$ and $T_R > 20$ are indistinguishable from ordinal preference then any theory--ratio, interval or ordinal--should give similar predictions or at best only as good as ordinal based predictions. Furthermore, we would expect the predictive capability to be similar to that for ordinal applied to the general population. Empirically for the 17% of the consumers with $T_o, T_R > 20$, we found that the predictions of each theory were almost identical giving an overall preference recovery of 57.1%, 28.6%, 14.3%, 0%, and 0%. This recovery agrees well with the ordinal theory (row 1 of table 3).

Thus the empirical evidence supports the conjecture that the intensity of preference theories can provide improved predictions and useful diagnostic information relative to existing theories such as conjoint analysis and preference regression. Existing techniques do well, as evidenced by tables 1 and 2, but the more general theory can make more efficient and effective use of consumer preference judgments. Clearly, this empirical example is a first application and test of the general theory but it does indicate that consumers can provide consistent CSPC judgments for products specified by attribute levels and that the majority of consumers give CSPC judgments that contain more extractable information than judgments based on the ordinal task used in conjoint analysis. Even for the 17% who gave at most ordinal results, the intensity measures did as well as the ordinal based conjoint analysis.

Summary

Consumer preference measurement is important for the design of new transportation and communications services. This paper has summarized the results of two improvements in the measurement of consumer preferences: (1) a general theory for intensity of preference which extends ordinal conjoint analysis to more efficient CSPC measures and (2) an interactive computer interviewing system and marketing research information system that makes this form of preference measurement more feasible.

The empirical results are promising. Direct comparisons show that the intensity theory as implemented through P.A.R.I.S. can improve existing techniques. Clearly, this is one test in a particular application, but it does raise important issues and is encouraging for further development of both the intensity measures and the interactive system.

Acknowledgment

The empirical study in section 4 was funded by a grant from the National Science Foundation, DS176-09479, Stuart Meyer principal investigator. We wish also to thank R. J. Taschek and T. Norris at Los Alamos Scientific Laboratory for their valuable assistance in the study.

References

- H. Abelman, "OPTISEP--A Package of Subroutines for Non-Linear Optimization," Vogelback Computing Center, Northwestern University, Evanston, Illinois, Document No. 322, July 1976.
- F. J. Carmone, Paul E. Green, and A. K. Jain, "The Robustness of Conjoint Analysis: Some Monte Carlo Results," Working Paper, University of Pennsylvania, May 1976.
- P. Cattin and D. R. Wittink, "A Comparison of Estimation Procedures for Multiattribute Models," Presented at the Joint National Meeting of the Operations Research Society of America and The Institute for Management Sciences, November 3-5, 1976, Miami, Florida.
- C. Cohen, "SUMT Version 4: Sequential Unconstrained Minimization Techniques for Non-Linear Programming," Vogelback Computing Center, Northwestern University, Evanston, Illinois, Document No. 200 (Rev. A), February 1975.
- P. H. Farquhar, "A Survey of Multiattribute Utility Theory and Applications," Management Science (forthcoming, 1977).
- P. E. Green, "On the Robustness of Multidimensional Scaling Techniques," Journal of Marketing Research, XII (February 1975), 73-81.
- P. E. Green and Y. Wind, "New Way to Measure Consumer's Judgments," Harvard Business Review (July-August 1975).
- J. R. Hauser, "Testing the Accuracy, Usefulness, and Significance of Probabilistic Choice Models: An Information Theoretic Approach," Working Paper, Northwestern University, Transportation Center, April 1976 (forthcoming, Operations Research).
- J. R. Hauser, "Consumer Preference Axioms: Behavioral Postulates for Describing and Predicting Stochastic Choice," Working Paper, Department of Marketing, Northwestern University, Evanston, Illinois, November 1976.

- J. R. Hauser and S. M. Shugan, "Efficient Measurement of Consumer Preference Functions: A General Theory for Intensity of Preference," Working Paper No. 602-001, Northwestern University, Department of Marketing, May 1977.
- J. R. Hauser and G. L. Urban, "A Normative Methodology for Modeling Consumer Response to Innovation," Operation Research, 25 (July-August 1977a).
- J. R. Hauser and G. L. Urban, "Direct Assessment of Consumer Utility Functions: von Neumann-Morgenstern Theory Applied to Marketing," Working Paper, MIT Sloan School, Revised January 1977b.
- R. M. Johnson, "Beyond Conjoint Measurement: A Method of Pairwise Tradeoff Analysis," Proceedings of the Association of Consumer Research (Cincinnati, Ohio: October 1975).
- R. M. Johnson, "Tradeoff Analysis of Consumer Values," Journal of Marketing Research, 11 (May 1974), 121-127.
- R. L. Keeney, "Multiplicative Utility Functions," Operations Research, 22 (January 1974), 22-23.
- R. L. Keeney and H. Raiffa, Decision Analysis with Multiple Conflicting Objectives (New York: John Wiley & Sons, 1976).
- Market Facts, Inc. and Peat, Marwich and Mitchell Co., "A Marketing Approach to Carpool Demand Analysis," Prepared for the Federal Energy Administration, Contract No. C-04-50179-00, April 1976, Washington, D.C.
- D. McFadden, "Conditional Logit Analysis of Qualitative Choice Behavior," in Paul Zarembka (ed.), Frontiers in Econometrics (New York: Academic Press, 1970), 105-142.
- S. H. Neslin, "Analyzing Consumer Response to Health Innovations: The Concept of Preference Inertia," Research Paper, MIT Sloan School of Management, Cambridge, Mass., May 1976.
- E. A. Pessemier, Product Management: Strategy and Organization (New York: Wiley-Hamilton, 1977).
- H. Raiffa, Decision Analysis Introductory Lectures on Choices Under Uncertainty (Reading, Mass.: Addison-Wesley, 1970).
- L. S. Shapley, "Cardinal Utility from Intensity Comparisons," RAND Report R-1683-PR, July 1975, Santa Monica, California
- S. M. Shugan and J. R. Hauser, "P.A.R.I.S.: An Interactive Market Research Information System," Working Paper, Northwestern University, Department of Marketing, Evanston, Illinois, May 1977.
- A. J. Silk and G. L. Urban, "Pretest Market Evaluation of New Packaged Goods: A Model and Measurement Methodology," Working Paper, MIT Sloan School of Management, Cambridge, Mass., February 1976 (forthcoming, Journal of Marketing Research, 1977).
- V. Srinivasan and A. Shocker, "Linear Programming Techniques for Multidimensional Analysis of Preferences," Psychometrika, 38 (September 1973), 337-370.
- H. M. Ting, "Aggregation of Attributes for Multi-attributed Utility Assessment," Technical Report No. 66, Operations Research Center, MIT, Cambridge, Mass., August 1971.
- W. S. Torgenson, Methods of Scaling (New York: John Wiley & Sons, 1958), 105-112.
- A. Tversky, "A General Theory of Polynomial Conjoint Measurement," Journal of Mathematical Psychology, 4 (1967), 1-20.
- G. L. Urban, "PERCEPTOR: A Model for Product Positioning," Management Science, 8 (April 1975), 858-871.
- J. von Neumann and O. Morgenstern, The Theory of Games and Economic Behavior, 2nd ed. (Princeton, N.J.: Princeton University Press, 1947).

CONSUMER RESEARCH IMPLICATIONS OF RANDOM COEFFICIENT MODELS

N. Scott Cardell
Ricardo Dobson
Frederick C. Dunbar
Charles River Associates
1050 Massachusetts Avenue
Cambridge, Massachusetts 02138

Abstract

Random coefficient models, in general, and the Hedonic Demand Model (HDM) in particular, are reviewed to show their basic properties and relevance to consumer research. Links to the three traditional tools of multiattribute models, individual differences scaling, and market segmentation are established. An empirical application of the HDM to auto demand is described. The conclusion integrates the review of the HDM and random coefficient models, generally, with consumer research objectives and requirements.

Introduction

Recent econometric modeling developments using random coefficients for probability choice behavior can be linked to prior consumer research on multiattribute models, individual differences scaling, and market segmentation. An experimental econometric technique estimates a hedonic demand model which incorporates selected elements from each of the other three standard analytical procedures from market research. Therefore, the new econometric modeling developments can be readily inserted into the array of consumer research analysis tools by reviewing the standard methodological procedures to show their relationship to the hedonic demand model.

Consumer research procedures, such as multiattribute models, are often connected with psychological assumptions of one kind or another. For example, Fishbein (1967) speculated that attitudes toward an object are a function of beliefs about the attributes of that object and the value of those attributes to a consumer. Several variations of his basic assumption have been proposed by Lehmann (1971), Hansen (1969) and others, and empirical research evaluates the nature of the relationships between attitudes toward an object, beliefs about the attributes of a product, and the importance or value of those attributes to consumers (Sheth and Tuncalp, 1975; Sheth, 1973). Individual differences scaling is another consumer research tool which relies heavily on psychological principles. In this case, the early work on nonmetric multidimensional scaling (Shepard, 1962a, 1962b; Kruskal, 1964a, 1964b) was followed into its extension so that each individual could be represented with a separate multidimensional space (Carroll, 1972; Carroll and Chang, 1970; Cliff, 1968; Tucker and Messick, 1963). Applications of individual differences scaling within consumer research are multitudinous (e.g., Green and Rao, 1972; Dobson and Nicolaidis, 1976; Benjamin, 1977).

Market segmentation builds on sociological concepts which respect the importance of groups of consumers relative to individual consumers. A fundamental tenet of market segmentation is that people have different preferences for products and services, but no individual, at least not a representative one, is totally distinct. Numerous empirical applications of market segmentation have been reported (Bass, Tigert, and Lonsdale, 1968; Blattberg and Sen, 1974; Haley, 1968; Johnson, 1971).

The hedonic demand model (HDM) falls in the tradition of econometric random coefficient models. It is related to

multiattribute research because it presumes that the utility of a product is based on one or more of its attributes. Since different individuals are allowed to differentially weight these attributes in the assessment of a product's utility, HDM shares features with individual differences scaling. Furthermore, the model is designed to work with aggregate databases, such as those which report only market shares, and therefore the orientation of HDM is compatible with market segmentation.

This presentation has three principal sections. Initially, a theoretical overview of random coefficient models is provided. The second section builds on the overview to establish a bridge between HDM and standard consumer research tools. This conceptual bridge should facilitate the adoption of HDM in consumer research. Finally, the third section describes some empirical results derived by applying the modeling orientation to the demand for autos.

Theoretical Overview

Random coefficient models can be characterized by their concern with the aggregate behavior of consumers although assumptions about individual differences are inherent to them. The models concentrate on *how many* consumers would buy a good rather than on *which* consumers would buy it. Traditionally, such models have been intrinsically linear and this limited their application to a wide class of consumer products. More recently, random coefficient models have been applied to the problem of consumers choosing only one among a finite number of items. This modeling orientation builds on economic antecedents which represent the utility of goods and services in terms of a bundle of attributes (Court, 1941; Lancaster, 1966). Empirical implications of this modeling orientation have been explored in various empirical studies (see, for example, Griliches, 1972; Charles River Associates, 1977; Hausman and Wise, 1976).

The Linear Demand Model

The value of the random coefficient approach can be shown by considering its application to linear equations. The demand relationships in linear form can be written as follows:

$$D_i = \sum_k \sum_j \beta_{kj} h_{kj} + \epsilon_i \quad (1)$$

where:

- D_i = the amount of the i th good to be consumed in aggregate,
- h_{kj} = the k th attribute of the j th good including its price,
- β_{kj} = estimated coefficients, and
- ϵ_i = independently distributed stochastic error term with mean zero and constant variance.

The above relationship, modified with demographic variables, is the workhorse of econometric demand studies with aggregate data.

There are several drawbacks to Equation (1). The first is that the coefficients will vary among individuals, thereby leading to biased estimates of the β_{kj} 's if they are assumed to be constant across the population (Theil, 1971). If a product's prices or attributes are changed, then the model will misforecast the change in demand.

To cope with this problem models have been developed which allow the β_{kj} 's to vary across individuals (Zellner, 1966; Swamy, 1974). It may be presumed, for example, that tastes for product attributes are distributed normally. If each β_{kj} is stochastically independent and if the h_{kj} 's are nonstochastic with finite variance and independent of the β_{kj} 's, then the mean and variance of the β_{kj} 's can be estimated with generally available regression methods. Theil (1971) points out that these restrictions are not necessarily innocuous. However, the model is appropriate to a large number of situations where consumer behavior needs to be inferred from aggregate data.

Probability Choice

Suppose that a product is such that consumers have a finite set of options and buy one such item on occasion to the exclusion of others. For example, a household may consider several different makes of autos, but it will usually purchase only one at a time. The time between purchases is usually substantial (e.g., more than a year). Other durable goods, such as refrigerators and televisions, are acquired in a similar fashion.

Now consider the problem of estimating the shares of various brands or models of such a product. It is well known that the linear specification given in Equation (1) misspecifies consumer behavior even if disaggregate data are used (Domencich and McFadden, 1975; Charles River Associates, 1973).

A class of models used to analyze consumer behavior under these conditions is based on the tenets of probability choice (Luce, 1959). They are widely used by transportation researchers (see Stopher and Meyburg, 1976, for a historical development of probability choice models in transportation). The most typical application is the use of the binary or multinomial logit model to estimate the probability of a traveler selecting one of several transport modes to make a trip. From a consumer research perspective, this decision can be generalized to the selection of one product or service over another. The probability of a consumer choosing alternative i (P_i) from n elements can be expressed as:

$$P_i = \frac{\exp(\beta_{ki} h_{ki})}{\sum_{j=1}^n \exp(\beta_{kj} h_{kj})} \quad (2)$$

where:

h_{kj} = the attributes (which are denoted by k) of each of the n alternatives, and
 β_{kj} = weights assigned to the attributes.

Equation (2) is a multinomial logit model, and it describes the probability of choice for alternative i as an S-shaped function of each of the product or service attributes. The sum of the P_i over individuals or households yields the demand, D_i , for the product.

Individual differences are not built into the underlying assumptions of the multinomial logit model. Sociodemographic descriptions of consumers can be inserted into Equation (2) to account for individual differences in alternative selection for cases in which the attributes of alternatives are invariant.

The probability choice framework is an advance but the assumption of constant coefficients is very limiting. To see this, suppose everyone in the population can choose among the same set of products. In a cross section of data, this means that the h_{kj} are constant for everyone. Thus, they have no explanatory power for observed differences in choices among products. Even if we have a time series of aggregate observations where changes in product shares can be correlated with changes in product attributes, we cannot estimate the β_{kj} 's in Equation (2) using standard statistical techniques (see McFadden, 1974b).

Even with the important theoretical advance of probability choice models, the problem of predicting product acceptance cannot be easily solved using aggregate data and assuming constant coefficients in the attributes of products. There are other options such as using special surveys or experiments, both of which may be costly. Alternatively, one can drop the assumption of constant coefficients and use widely available aggregate data. This option is explored below.

The Hedonic Demand Model

The hedonic demand model (HDM) allows using aggregate data to estimate market shares among very similar products. It is based on the theory of probability choice behavior while taking account of differences in individuals' tastes. The following develops the behavioral foundations of the model.

A basic assumption is that consumers select that alternative which yields the greatest utility. The utility (denoted by U) derived by a consumer (denoted by x) confronted with several discrete goods, D_i , and all other goods and services, Z , can be represented by:

$$U(D_i, Z|x). \quad (3)$$

It is useful to require that expenditures on Z and the selected alternative equal a consumer's income, I_x , so that:

$$I_x = D_i P_i + \pi Z, \quad (4)$$

where:

$D_i = 1$ for the selected alternative and 0 for all other alternatives,
 P_i = alternative prices, and

πZ = the price index for all other goods and services.

Therefore, the choice situation being modeled can be represented by:

$$U(D_j, Z|x) \geq U(D_i, Z|x) \text{ for all } D_i, i \neq j, \quad (5)$$

and the budget constraint of

$$I_x = D_j p_j + \pi Z. \quad (6)$$

It is also convenient to assume separability between utility based on D and Z so that the link to product attributes can be shown as:

$$U(h_j|x) \geq U(h_i|x) \text{ for all } i \neq j \quad (7)$$

where:

h_j = the vector of attributes of good j , and

x = the vector of taste characteristics.

Since Z does not vary when different D_i are evaluated, Equation (7) is equivalent to Equation (5) under the assumption of separability.

A further assumption of separability and additivity among the attributes makes it possible to define

$$U(h_j|x) = \sum_{k=1}^m h_{kj} x_k \quad (8)$$

where:

h_{kj} = attributes, such as size, horsepower, and price in the case of an auto, and

x_k = the tastes of a consumer with respect to those attributes.

There are m attributes under consideration.

Different consumers are represented by different x_k . The distribution of tastes across consumers and over attributes can be characterized by multivariate probability density functions.

It is useful for interpretive purposes to normalize the x_k so that the taste coefficient of price for each consumer is -1. This normalization permits Equation (8) to be rewritten as:

$$U(h_j|x) = \sum_{k=1}^m h_{kj} x_k - p_j \quad (9)$$

where:

x_k = marginal utilities which indicate the price a consumer would pay for another unit of the k th attribute.

Even when one product dominates another on all attributes, it might be selected by a small fraction of consumers. In order to simulate this potential occurrence, it is convenient to allow a stochastic error term to be added to the utility function of Equation (9). It is usually assumed that the error term is distributed independently and identically across consumers. With the error term, the utility of alternative j can be represented as:

$$U(h_j|x) = \sum_{k=1}^m h_{kj} x_k - p_j + \epsilon_j \quad (10)$$

As a consequence of this error term, the probability of j being selected over i can be denoted by:

$$\text{prob}[U(h_j) > U(h_i)] \quad (11)$$

$$= \text{prob} \left[\sum_{k=1}^m x_k (h_{kj} - h_{ki}) + p_i - p_j > \epsilon_i - \epsilon_j \right] \quad (12)$$

Therefore, for the fraction of the cases in which the $\epsilon_i - \epsilon_j$ is greater than the quantity on the left of Equation (12), alternative i will be selected even when it is dominated on all attributes by alternative j .

While the ϵ 's are distributed independently according to a scaled Weibull distribution, at the level of individual consumers we have a multinomial logit model. However, the parameters of HDM differ across individuals. It is not appropriate, therefore, to apply the multinomial model directly to aggregate market share data. It is necessary instead to estimate the moments (i.e., the mean, variance, and higher ones if appropriate) of an assumed distribution of taste differences across individuals. As the random coefficient models are normally applied, the distribution of taste differences is not directly observed. The distribution is assumed to follow a suitable density function, such as log normal. The data requirements of HDM are substantially reduced by working with an assumed instead of an observed distribution of taste parameters.

McFadden (1974a) has shown that in this case the probability of an individual choosing good j from among n goods given tastes x_k is:

$$P_j = \frac{e^{[U(h_j|x_k)]}}{\sum_{i=1}^n e^{[U(h_i|x_k)]}} \quad (13)$$

Equation (13) can be used to help derive the mean, μ , and the standard deviation, σ , of consumer tastes by integrating over x_k :

$$P_j(\mu, \sigma) = \int \dots \int \frac{e^{[U(h_j|x_k)]}}{\sum_{i=1}^n e^{[U(h_i|x_k)]}} f(x_k; \mu, \Sigma) dx_k \quad (14)$$

where $f(x_k; \mu, \Sigma)$ is the probability density function for the multivariate distribution of consumer tastes over attributes with mean vector μ and variance-covariance matrix:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_m^2 \end{bmatrix} \quad (15)$$

The probability expression of Equation (14) involves a rather complex integral -- its order of integration is the same as the number of attributes. While it is impossible to analytically evaluate Equation (14) except under severely simplifying assumptions, numerical integration is possible. Numerical integration relies on Monte Carlo procedures. $P_j(\mu, \sigma)$ can be approximated by taking the average of many calculated P_j from Equation (14) where x_k is a random vector generated from a multivariate probability density function with mean μ and variance-covariance matrix Σ . The details of the solution algorithm and the specification of functional forms of x_k are discussed at length elsewhere (Charles River Associates, 1977) and they will not be summarized here since they are not essential to the points being made.

Ties to Standard Consumer Research Techniques

Random coefficient models are important for consumer research because of the wide range of topics to which they are relevant and the inherent power of the research strategy. In this section, we describe relationships between standard consumer research techniques and HDM in order to highlight the value of this recent econometric development vis-a-vis the analysis of consumer behavior. In particular, three techniques are compared and contrasted with HDM: multiattribute models; individual differences scaling; and market segmentation.

Multiattribute Models

A common form of the multiattribute model has been described by Wilkie and Pessemier (1973). They write a linear compensatory representation in these symbols:

$$A_{jk} = \sum_{i=1}^n I_{ik} B_{ijk} \quad (16)$$

where

i = attribute or product characteristic,

j = brand, and

k = consumer or respondent;

such that:

A_{jk} = consumer k 's attitude score for brand j ,

I_{ik} = the importance weight given to attribute i by consumer k , and
 B_{ijk} = consumer k 's belief as to the extent to which attribute i is offered by brand j .

Consumer researchers have extensively studied variations of this formulation. The original attitude conceptualizations were proposed by Rosenberg (1956) and Fishbein (1967). The relationships among A_{ijk} , I_{ik} , and B_{ijk} have been examined under different measurement conditions and modeling frameworks.

A gap in the theoretical development of early multiattribute formulations was the nature of the relationship between A_{ijk} and consumer behavior. The link between attitude and behavior is implicit with positive attitudes assumed to correlate with positive behavior, but it is known that better predictors of behavior are available.

The strength of the approach is diagnosis as opposed to prediction. Diagnosis results from the exposure of the salience of different product attributes as seen by consumers. To the extent that managers can change product attributes, they can use multiattribute modeling research to help them improve the salience of their product to consumers.

The utility equation (see Equation (10)) of HDM is defined in terms of coefficients for different attributes, in a similar manner to multiattribute models. However, the attributes, h_{kj} , are direct observations on products as opposed to respondents' judgments about product attributes. The utility concept is directly linked to consumer choice through utility maximization and probability representations (see Equations (13) and (14)). Finally, the taste coefficients, x_k in Equation (10) generally correspond to I_{ik} of Equation (16), but they are estimated from aggregate data on market shares instead of being reported by respondents. Therefore, the random coefficient models do not require data from individual respondents. In the latter sense, they are more parsimonious in their data needs; this reduces the cost of analysis and thus adds to the desirability of HDM.

The theoretical basis of HDM is inherently reasonable, and its relationship to consumer behavior is explicit. Consumers are assumed to select the product with the largest utility to them, and utility is defined as a weighted function of attributes and a stochastic term. The weights for attributes are consumer taste coefficients which are estimated to maximize the likelihood of predicting aggregate market share data across respondents.

Individual Differences Scaling

Individual differences scaling follows from psychological research on perceptions and preferences. The notion is that the differences among individuals in preferences or perceptions can be used to quantify the similarities and differences among individuals. With respect to marketing research, it is assumed that these differences have implications with respect to consumer behavior. Among the models which have been commonly used for individual differences scaling are INDSCAL (Carroll and Chang, 1970), the vector model (Tucker, 1960), and variations of the ideal point model (Coombs, 1964).

There are at least four types of applications for individual differences scaling in consumer research. In common with aggregate multidimensional scaling, the technique finds how product attributes vary among one another from the user's perspective. Beyond this point, individual differences scaling is much more flexible. It is possible, for example, to stratify a sample along one or more sociodemographic variables to study how shifts in perceptions and/or preferences vary across

strata. It is alternatively possible to cluster consumers according to their perceptions and/or preferences and examine shifts in sociodemographic variables over perceptual market segments. When it is more useful to find homogeneous perceptual segments than study variation along one or more prespecified sociodemographic variables, the latter procedure is to be preferred over the former. Both techniques, however, will reveal interrelationships of particular types of viewpoints with sociodemographic variables. A fourth application involves using an individual's weights or saliences for perceptual dimensions as the independent variables in a multiattribute formulation. The dependent variable is usually some indicator of consumer preference or purchase patterns. Illustrations of the above four applications are described in numerous sources (see for example, Dobson, Golob and Gustafson, 1974; Dobson and Kehoe, 1975; Green and Carmone, 1972; Green and Rao, 1972; and Dobson and Nicolaidis, 1976).

Individual differences scaling and HDM have complementary and contrasting relationships to each other. There are no guidelines within HDM to select attributes, but individual differences scaling can be used to highlight product attributes which are salient to major market segments. Physical indicators of these psychologically salient attributes could then be used in HDM. It is possible to use output from individual differences scaling in a multiple attribute representation which parallels that of HDM, but the interpretations of the two representations are substantially at variance. In the case of individual differences scaling, consumer tastes, as denoted by saliences for perceptual dimensions, are the independent variables. These tastes are then weighted to account for indicators of consumer preferences or purchase patterns. HDM, however, uses consumer tastes as weights instead of independent variables. Further, these consumer tastes are linked to behavior via a nonlinear, probability choice formulation rather than the linear, compensatory formulation frequently used with individual differences scaling output.

Both random coefficient models and individual differences scaling relate to market segmentation. The random coefficient models are attractive, in part because they can be estimated on aggregated market segmentation data. For HDM, the aggregate data are market shares for different products, a type of information which is frequently readily available. On the other hand, individual differences scaling will normally require a special data collection effort, but its output can be used to form market segments based on perceptual characteristics of respondents. For product positioning with respect to consumer segments in a broad, qualitative sense, the richness of individual differences scaling may be more useful, but for detailed, quantitative estimates of market shares, HDM may be more useful. In any event, there are opportunities for meaningful consumer research in comparing and contrasting the alternative methodologies.

Market Segmentation

Market segmentation in consumer research is predicated on two key concepts. It is initially assumed that meaningful groups of users of a product or service can be identified. The second concept revolves around the creative assessment of product and/or service requirements for different market segments. These concepts are jointly applied towards maximizing profits through rendering services and/or products which appeal strongly to lucrative market segments.

A variety of criteria have been proposed for assessing whether any particular group of market segments is appropriate for profit maximization. Kotler (1972) notes that it is important for segments to be measurable,

accessible, and substantial. Wilkie and Cohen (1976) remark that "true" segments should be homogeneous within and heterogeneous between groups, useful as a behavioral correlate, and efficient to market from production, distribution, and pricing perspectives. A pivotal aspect of market segmentation is the selection of appropriate variables on which to divide consumers. The term segmentation basis is used to denote a set of variables selected to divide consumers. Three principal classes of bases are sociodemographic data (Carman, 1965; Assael and Roscoe, 1976), consumption patterns (Cunningham, 1955), and individuals' subjective judgments (Haley, 1968; Myers, 1976). A major summary and analysis of segmentation research and theory is provided by Frank, et al. (Frank, Massey and Wind, 1975).

As noted above, HDM is organized around a segmentation structure, namely the market shares for different products. The primary thrust of this modeling orientation is to derive estimates of consumer tastes which are consonant with market share data. This consumer taste information can be used for product planning purposes. Also, market simulations can be performed which reveal the implications of an altered product mix. HDM assumes that the set of salient product attributes are known and that the purpose of a marketing analysis is to study substitution among products. When either of these conditions does not hold, another market segmentation analysis technique may be more appropriate. More traditional consumer research tools for market segmentation are directed at product positioning (i.e., what new product would consumers most like to have) and advertising targeting (i.e., what content should a promotional campaign have and to whom should it be addressed). A potentially useful purpose of the traditional approaches is the evaluation of marketing strategy effectiveness (Assael and Roscoe, 1976). It can be seen, therefore, that traditional consumer techniques and HDM complement each other with respect to market segmentation.

An Application of HDM to Autos

The hedonic demand model was developed in a study of the impact of foreign trade controls on the import share of new car sales (Charles River Associates, 1976). The technique yields predictions of the effects of changes in prices and characteristics of any set of automobile models on the market shares of individual models. The predictions of share changes for individual models are then summed to obtain the predicted change in the share of a class of models resulting from any relative price changes.

The input data used are the price, selected attributes, and new car sales in a given time period of a set of automobile models. The attributes include available characteristics which either contribute to, or are correlated with, the quality of a car. Available attributes data for recent years include both the physical dimensions from Automotive News Almanac and performance data collected by the U.S. Department of Transportation and the U.S. Environmental Protection Agency. The physical attributes data include number of cylinders, headroom, height, horsepower, legroom, length, gas tank capacity, transmission type (automatic or manual), turning circle weight and width. The performance variables used were acceleration and fuel economy (miles per gallon). The sample period was the five-month period between April 1974 and August 1974; this period was selected to include the part of the 1974 model year subsequent to the period of acute gasoline shortage.

The results presented below should be viewed as first-run estimates only. Still, the results seem quite plausible and are very similar both to estimates from time series regressions and to a set of predictions of

the impact of imported car price increases on market shares derived by questionnaire techniques (Market Facts, 1976).

The estimate of the taste distribution for characteristics which was used to derive the share elasticities with respect to relative price changes was based on five attribute variables: volume, passenger area, weight, turning circle, and gallons per mile. Volume is defined as the product of length, width, and height, and is a proxy measure for a combination of total seating capacity and trunk space. Passenger area is computed as the product of headroom and legroom, and represents a proxy for driver and front seat passenger comfort. Utility is positively correlated with volume, passenger area, and weight. Utility is negatively correlated with turning circle diameter (the smaller the turning circle, the more maneuverable the automobile) and with fuel consumption in gallons per mile.

Table 1 summarizes statistical information on the estimated distributions of the marginal rates of substitution between characteristics and price. The distribution of consumer tastes for auto attributes is assumed to be log-normal. The statistics reported in Table 1 were computed by taking exponents of the parameters of the normal distributions of the logs of the attributes. As the distributions are log-normal, the mean and median of the distributions are not the same.

TABLE 1
DISTRIBUTION OF ESTIMATED MARGINAL VALUES OF
AUTOMOBILE CHARACTERISTICS
(DOLLARS PER UNIT)

Characteristic	Median	Mean	Variance in Consumer Tastes	95 Percent Consumer Taste Range ¹	
				Absolute Lower Bound ²	Absolute Upper Bound ³
				(Col. 1)	(Col. 2)
VOL 2	2.3053	7.1149	431.57	.12156	43.717
AREAG	5.1795	5.8383	9.2219	1.9850	13.515
WEIGHT	2.7782	3.1316	2.6532	1.0648	7.2492
TURN	-7.4435	-303.29	152,630,000	-.03581	-1,547.4
GPM	-565.42	-637.33	109,990	-216.70	-1,475.30

Variable Definitions

VOL 2 = (.0001) x (LENGTH) x (WIDTH) x (HEIGHT) (in cubic inches)
 AREAG = (HEADROOM) x (LEGROOM) (in square inches)
 WEIGHT = weight (in lbs.)
 TURN = turning circle diameter (in feet)
 GPM = gallons per 100 miles

¹Range of values of marginal rates of substitution including 95 percent of consumers.

²2.5th percentile of consumer taste distribution.

³97.5th percentile of consumer taste distribution.

The numbers reported in Column (2) of Table 1, the median of the distributions, give the marginal value of an additional unit of the attribute in dollars; the units of the characteristics are defined at the bottom of Table 1. The coefficient attached to VOL 2 indicates that the median consumer is willing to pay \$2.31 for an extra 10,000 cubic inches of volume. The coefficient attached to weight indicates that the median consumer is willing to pay an additional \$2.78 for an additional pound. The median coefficient for turning radius is greater than for VOL 2, AREAG, and WEIGHT, but there is also the most differentiation among consumers on this attribute. Note the extremely large variance of consumer tastes for this attribute in Column (3). The median marginal utility for gallons per mile is the greatest of all the attributes. This coefficient, which was derived from data collected immediately after the oil embargo, shows the extreme sensitivity of consumers to gasoline availability. However, despite the time period for which the model is estimated, there is still substantial variability among consumers on the importance of this attribute.

The upper and lower bounds of the 95 percent confidence interval (see Columns (4) and (5)) indicate the degree of dispersion in tastes for each attribute among consumers. The alternative degrees of variability across attributes reveals the ability of HDM to characterize individual differences among consumers. It should also be kept in mind that the model was estimated on physical attributes data and aggregate market share statistics. In any event, it can be seen that the most widely dispersed values are for the attribute "turning circle." Apparently some consumers place a high value on this attribute for parking and general maneuvering ability, but others have practically no preference for it at all. It is likely that a substantial portion of the variation for gallons per mile reflected different expectations among consumers as to gasoline availability and price. The range of variability for the remaining three attributes is substantially narrower than for TURN and GPM.

Table 2 compares the predicted shares consistent with the utility function estimated in the model with actual market shares of different automobile types. The model was estimated using two different price variables -- the one year old price in August 1974 of 1974 models reported in National Market Reports, Inc., Red Book and the estimated new transactions price of 1974 models. The new transactions price variable was corrected by applying "typical" discount factors for different types of cars reported in the annual automobile buying issue of Consumer Reports (April 1974) to list prices reported in the 1974 issue of Automotive News Almanac. The data reported in Table 2 show that the predicted shares from the model conform fairly closely to the actual shares; the model fits the data well. For example, the actual share for all domestics was .86798 and the share estimated by the random coefficient model with adjusted new car prices was .86183. Predicted import shares are slightly higher than actual shares. There is little difference in model estimates across the two price indicators which suggests the model is robust in the face of varying data quality.

TABLE 2
COMPARISON OF PREDICTED SHARES TO ACTUAL SHARES
BY AUTOMOBILE TYPE, APRIL-AUGUST 1974

	Actual Share	Predicted Share: One Year Old Price	Predicted Share: Estimated New Transactions Price
Domestic Autos			
Subcompact	.08792	.07900	.07767
Compact	.21413	.23299	.23701
Luxury Small	.03397	.01825	.01822
Intermediate	.26943	.24582	.25633
Standard	.20846	.21383	.20902
Luxury Standard	.05108	.05598	.06075
Specialty	.00299	.00215	.00223
All Domestics	.86798	.84802	.86183
Imported Autos			
Subcompact	.08013	.09624	.09080
Compact	.05340	.04449	.03744
Sports Car	.01200	.00815	.00692
Luxury	.00444	.00321	.00336
All Imports	.13197	.15209	.13852

Summary and Conclusions

A major purpose of this report was to present a recent econometric development which has potential value as a consumer research tool. The historical antecedents and underlying assumptions were initially described. These assumptions were subsequently compared and contrasted with those of three standard consumer research techniques. Finally, an application of the approach to auto demand was summarized. It is clear that HDM parallels the functions of other consumer research techniques, that it requires little data which can be obtained without survey collection costs and time delays, and that

it has the potential to produce accurate results which are subject to rich interpretation.

The logic of HDM is based on a simple assumption of utility maximization at the level of individual consumers. Furthermore, the formulation of this model assumes a multiattribute structure, but it is based on physical product attributes. Individual differences among consumers are represented by taste coefficients for the attributes. These coefficients are estimated in a probability choice framework to be consistent with market share data.

As mentioned above, HDM is multiattribute in structure and therefore has a common link with the Fishbein (1967) and Rosenberg (1956) models. A principal difference between the two approaches is the use of physical attribute data versus subjective judgments in the psychological models. The individual differences scaling models, such as INDSICAL, allow for variations among consumers in a fashion which is similar to HDM. However, neither individual differences scaling nor Fishbein models, as initially proposed, were designed to facilitate the estimation of the probability of consumer choice, but HDM is directed toward this objective. Market share data, a type of market segmentation, provide a basis for the estimation of HDM. In the latter respect, a tie is established to a pervasive consumer research tool.

When the model was estimated on auto market share data, its properties were revealed. It was shown that variability among consumers could be represented. It is important to recall that this variability among consumers is described without a reliance on surveys. Only readily available data, such as market shares for models and physical product attributes, were required. In addition, it was possible to estimate a model representation which accurately fitted the data. This ability to accurately describe variability among consumers while correctly estimating market share data is an attractive property, particularly when it is combined with the minimal data requirements of the random coefficient models.

Further research is justified with random coefficient models on the basis of the results described above. These models can have rational structures which show a relationship between utility and probability choice at the level of the individual consumer. They also have extremely simple data requirements. HDM is compatible with standard consumer research functions. Therefore, its use in marketing studies will achieve normal objectives and concurrently extend the state of the art in consumer research.

References

- H. Assael and A. M. Roscoe, "Approaches to Market Segmentation Analysis," Journal of Marketing, 40 (1976), 67-76.
- F. M. Bass, D. G. Tigert, and R. F. Lonsdale, "Marketing Segmentation: Group Versus Individual Behavior," Journal of Marketing Research, 5 (1968), 264-270.
- J. M. Benjamin, The Assessment of User Preferences for New Apartments: An Application of Multidimensional Scaling. (unpublished doctoral dissertation, State University of New York, Buffalo, 1977).
- R. C. Blattberg and S. K. Sen, "Market Segmentation Using Models of Multidimensional Purchasing Behavior," Journal of Marketing, 38 (1974), 17-28.
- J. M. Carman, Studies in Demand for Household Equipment, Research Program in Marketing, University of California, Berkeley, 1965.

- J. Carroll, "Individual Differences in Multidimensional Scaling," in R. N. Shepard, A. K. Romney, and S. B. Nerlove, eds., Multidimensional Scaling I (New York: Seminar Press, 1972).
- _____ and J. J. Chang, "An Analysis of Individual Differences in Multidimensional Scaling via an N-way Generalization of 'Eckart-Young' Decomposition," Psychometrika, 35 (1970), 283-319.
- Charles River Associates, Analysis of the Demand for Cable Television (Cambridge, Massachusetts, 1973).
- _____, Impacts of Trade Policies on the U.S. Automobile Industry (Washington, D.C.: U.S. Department of Labor, 1976).
- _____, Methodology for Predicting the Demand for New Electricity Using Goods (Palo Alto, California: Electric Power Research Institute, 1977).
- N. Cliff, "The 'Idealized Individual' Interpretation of Individual Differences in Multidimensional Scaling," Psychometrika, 33 (1968), 225-232.
- C. H. Coombs, A Theory of Data (New York: John Wiley and Sons, Inc., 1964).
- L. M. Court, "Entrepreneurial and Consumer Demand Theories for Commodity Spectron," Econometrica, 1941.
- R. Cunningham, "Measurement of Brand Loyalty," Proceedings of the American Marketing Association Conferences, 1955, 39-45.
- R. Dobson, T. F. Golob, and R. L. Gustafson, "Multidimensional Scaling of Consumer Preferences for a Public Transportation System: An Application of Two Approaches," Socio-Economic Planning Sciences, 8 (1974), 23-36.
- _____ and J. Kehoe, "An Empirical Comparison of the Tucker-Messick and INDSCAL Models: Measuring Viewpoints about Transit Attributes," Socio-Economic Planning Sciences, 9 (1975), 67-81.
- _____ and G. C. Nicolaidis, "Preferences for Transit Service for Homogeneous Groups of Individuals," Proceedings, Fifteenth Annual Meeting, Transportation Research Forum, 1976.
- T. Domencich and D. McFadden, Urban Travel Demand (Amsterdam: North-Holland Publishing Co., 1975).
- M. Fishbein, "A Behavior Theory Approach to the Relations between Beliefs about an Object and the Attitude toward an Object," in M. Fishbein, ed., Readings in Attitude Theory and Measurement (New York: John Wiley and Sons, Inc., 1967).
- R. Frank, W. F. Massy, and F. Wind, Market Segmentation (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1975).
- P. E. Green, and F. J. Carmone, "Marketing Research Applications of Non-metric Scaling Methods," in A. K. Romney, R. N. Shepard, and S. B. Nerlove, eds., Multidimensional Scaling II (New York: Seminar Press, 1972).
- _____ and V. R. Rao, Applied Multidimensional Scaling (New York: Holt, Rinehard, and Winston, Inc., 1972).
- Z. Griliches, ed., Price Indexes and Quality Change (Cambridge, Massachusetts: Harvard University Press, 1972).
- R. I. Haley, "Benefit Segmentation: A Decision Oriented Research Tool," Journal of Marketing, 37 (1968), 30-35.
- F. Hansen, "Consumer Choice Behavior: An Experimental Approach," Journal of Marketing Research, 6 (1969), 436-443.
- J. Hausman and D. Wise, "A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences," Working Paper, Department of Economics, Harvard University, 1976.
- R. M. Johnson, "Market Segmentation: A Strategic Management Tool," Journal of Marketing Research, 8 (1971), 13-18.
- P. Kotler, Marketing Management (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1972).
- J. B. Kruskal, "Multidimensional Scaling by Optimizing to a Nonmetric Hypothesis," Psychometrika, 39 (1964), 1-27.
- _____, "Nonmetric Multidimensional Scaling: A Numerical Method," Psychometrika, 29 (1964), 115-129.
- K. J. Lancaster, "A New Approach to Consumer Theory," Journal of Political Economy, 74 (1966), 132-157.
- D. R. Lehmann, "Television Show Preference: Application of a Choice Model," Journal of Marketing Research, 8 (1971), 47-55.
- R. Luce, Individual Choice Behavior (New York: John Wiley and Sons, Inc., 1959).
- Market Facts, A Foreign Automobile Impact Study (Washington, D.C.: U.S. Department of Labor, 1976).
- D. McFadden, "Conditional Logit Analysis of Qualitative Choice Behavior," in Paul Zarembka, ed., Frontiers in Econometrics (New York: Academic Press, 1974a).
- _____, "The Measurement of Urban Travel Demand," Journal of Public Economics (1974b).
- J. H. Myers, "Benefit-Structure Analysis: A New Tool for Product Planning," Journal of Marketing, 40 (1976), 23-32.
- J. M. Rosenberg, "Cognitive Structure and Attitudinal Affect," Journal of Abnormal and Social Psychology, 53 (1956), 367-372.
- R. N. Shepard, "The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function. I," Psychometrika, 27 (1962a) 125-140.
- _____, "The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function. II," Psychometrika, 27 (1962b) 219-246.
- J. N. Sheth, "Brand Profiles from Beliefs and Importances," Journal of Advertising, 13 (1973), 37-42.
- _____, and S. Tuncalp, Structural Assumptions Underlying Fishbein's Expectancy Value Model of Attitudes (Champaign, Illinois: University of Illinois, 1975).
- P. R. Stopher and A. H. Meyburg, "Behavioral Travel-Demand Models," in P. Stopher and A. H. Meyburg, eds., Behavioral Travel Demand Models (Lexington, Massachusetts: Lexington Books, 1976).

P. A. V. B. Swamy, "Linear Models with Random Coefficients," in P. Zarembka, ed., Frontiers in Econometrics (New York: Academic Press, 1974).

H. Theil, Principles of Econometrics (New York: John Wiley and Sons, Inc., 1971), Chapter 11.

L. R. Tucker, "Intra-individual and Inter-individual Multidimensionality," in H. Gulliksen and S. Messick, eds., Psychological Scaling: Theory and Applications (New York: John Wiley and Sons, Inc., 1960).

_____ and S. Messick, "An Individual Difference Model for Multidimensional Scaling," Psychometrika, 28 (1963), 333-367.

W. L. Wilkie, and J. B. Cohen, A Behavioral Science Look at Market Segmentation Research (Gainesville: University of Florida, 1976).

_____ and E. A. Pessemier, "Issues in Marketing's Use of Multi-Attribute Attitude Models," Journal of Marketing Research, 10 (1973), 428-441.

A. Zellner, "On the Aggregation Problem: A New Approach to a Troublesome Problem" Report #6628, Center for Mathematical Studies in Business and Economics, University of Chicago, 1966.

Attitude-Behavior Changes in A
Before-After Mode Choice Situation

Mary Lynn Tischer
Carl G. Shea
Urban Planning Division,
Federal Highway Administration
U.S. Department of Transportation

Abstract

A non-equivalent control group design was utilized in the evaluation of the effects of a high-occupancy reserved lane on the mode perceptions and mode choice behaviors of downtown commuters. Significant behavior and perception change was observed in both the treatment and control groups. The standardized change score analysis suggested that the transportation modification had some effect on perceptions although no effect on behavior was observed.

Introduction

A major objective of current transportation policy is the attainment of a more balanced transportation system through greater use of ridesharing modes. Programmatic efforts to facilitate the use of mass transit and carpooling must be based, however, on knowledge of the causal components of individual decisionmaking and an awareness of the level of transportation modifications necessary to affect behavior. Interviewing the same individuals both before and after transportation modifications have been effected can provide the necessary information regarding the relationships among system characteristics, individual perceptions, preferences and social situations and transportation behavior.

Panel data have previously been utilized in transportation research to evaluate the impacts of major changes in the transportation-related environment. Watson and Holland (1977) interviewed a Singapore sample twice to obtain behavioral information on the effects of an urban auto restraint policy. Skinner (1975) used repeated measurements of travel behavior to assess the effects of increased gasoline prices during the 1975-76 energy crisis. Longitudinal data have also provided information on the reliability of models in forecasting response to new modes (Parody, 1977; McFadden, et al., 1977) and on the relationship between attitudinal structures and behavior in response to transportation innovations (Goss and Shuldiner, 1977; Gilbert et al., 1975). Little information is available, however, on the effects of transportation system management actions (TSM) on the perceptions, preferences and behaviors of individuals.

The relationship of the system to the individual and his social environment is represented in Figure 1. While the figure does not detail a theory of behavior, it does provide a framework for the analysis of potential effects of alterations in the transportation system. It is assumed that changes in the system affect the service provided by the modes which, when communicated to individuals, alters their perceptions. Similarly, it is suggested that changes in personality variables, for example an increase in the value of a clean environment, will change the perceptions of the instrumentation necessary to reach the goal and lead to the development of new behaviors.

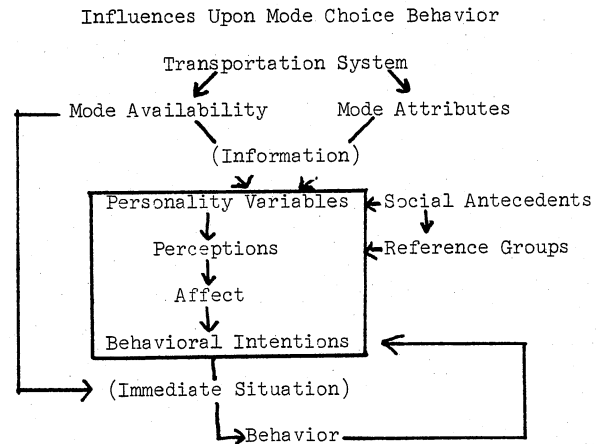
The objective of this paper is to present preliminary findings on the effects of a reserved lane for buses and carpools implemented on the Santa Monica, California, freeway on the mode perceptions and mode choice behaviors for downtown Los Angeles commuters. Changes in individuals residing in this corridor are compared to those observed for a sample of commuters residing

along comparable freeways which also feed into the central business district.

The Design and Methodology

The introduction of a control group is a necessary feature when one is interested in the effects of a transportation modification; otherwise, changes in behavior could also be attributed to extraneous factors operating in the environment. The effects of the modification, or "treatment," are distinguished from general system activities through the comparison of two similar groups, only one of which was exposed to the transportation system changes. In the analysis of social phenomena where similarity of individuals cannot be obtained through random assignment into treatment conditions, the experimental design can be approximated with a "non-equivalent control group design," (Campbell and Stanley, 1963). The similarity of control and treatment groups was obtained, in this case, by sampling within census tracts matched by socio-demographic and travel-related characteristics, (Hartgen and Howe, 1976).

Figure 1



There has been considerable debate as to the statistical technique appropriate for analyzing changes in the non-equivalent control group design. Selection of treatment group based on group differences can result in under-identification of the change model (Linn and Werts, 1977). However, if it is reasonable to assume that any group differences which might confound the treatment with the dependent variable are stationary over time, change score analysis can be applied (Kenny, 1975).

To perform the analysis of treatment effects, the variables are first standardized to correct for possible changes in variance. The time 1 treatment and control group scores for the variable of interest are combined and standardized; the time 2 scores are similarly adjusted. Differences between treatment and control groups are then evaluated with a correlational technique in which the treatment variable is dummy coded and the standardized pre-and-post test scores are correlated with the treatment variable. A significant difference between the pre-test-treatment correlation and the post-test-treatment correlation is indicative of a treatment effect. Significance is determined with the t test for dependent correlations (McNemar, 1969). The t test for related samples is used to evaluate the significance of changes within the treatment and control groups (Bruning and Kintz, 1968).

The Sample and Data

Census tracts within two miles of the Santa Monica Freeway in Los Angeles, California with a high incidence of Central Business District (CBD) workers were identified. These census tracts were matched by socio-demographic characteristics with tracts surrounding alternative freeways which also fed into downtown Los Angeles. Within the census tracts, households were randomly located through reverse telephone directories. Only individuals who worked within the business district were eligible for the home interview. Where a household contained multiple workers, the individual taking the lesser used mode was selected. A supplemental sample of carpoolers observed entering the CBD and bus users located at bus stops was included to ensure representation of these modal groups.

Interviews with 996 individuals were obtained between September 1, 1975 and March 5, 1976; 512 lived in the Santa Monica corridor and 484 were in the control areas. From January - May 1977, attempts were made to recontact the 996 individuals. Interviews were completed for 502 of the Phase I respondents.

Individuals were asked how frequently each month they drove to work alone, drove another person, rode with another and took a bus. Perceptions for each of 3 modes (sharing a ride, driving alone, and taking a bus) were collected for 19 attributes using a seven point semantic differential format. The perceptions were coded such that seven represents the most positive aspect of each attribute. The attribute labels are presented in Table 1.

Table 1
Attribute Labels

Label
1. Comfort
2. Convenience
3. Expense
4. Package space
5. Ease of use
6. Reliability
7. On time arrival
8. Traveltime
9. Vehicular safety
10. Personal safety
11. Destination accessibility
12. Crowding
13. Wait for vehicle
14. Relaxing
15. Weather
16. Waiting in traffic
17. Flexible schedule
18. Extra time
19. Parking cost

The same questions were asked at the two interviews. However, at time 2, individuals were asked to recollect their behavior and perceptions from the period of March through August 1976.

The Treatment

From March 15 to August 8, 1976, the median lanes on the Santa Monica Freeway were reserved for the exclusive use of buses and automobiles carrying three or more persons. No barriers separated the reserved lanes from general traffic, thereby permitting entry or exit throughout the 12.5 miles of the freeway. Preferential freeway access for vehicles carrying two or more persons had been implemented on 12 of 30 ramps prior to the Phase I interviewing although metering rates for the non-bypass ramp lanes were drastically altered at

the start of the diamond lane project. Four new express bus routes were established and four existing bus routes took advantage of the lane. Three new routes were developed to serve the three newly-opened park-and-ride lots and a crosstown feeder service was initiated. Prior to the project, 18 express bus trips were offered every weekday morning, while on the first weekday in August, 74 express buses were in operation. Bus headways were also improved; a policy maximum of 15 minute headways was established during peak hours. Bus prices were increased resulting in an average rise of 20¢ per trip (Billheimer, et al., 1977). On August 11, 1976, the preferential lanes on the Santa Monica Freeway were terminated by court order; improvements in the bus system remained in effect. The freeways in the control areas were not modified during this period, although bus service was improved throughout the Los Angeles region and bus fares were increased in the Southern California Rapid Transit District.

Sample Mortality, Control and Treatment Group Equivalence

Only 50% of the Phase I respondents were reinterviewed at time 2. Thirteen percent of the initial sample refused a second interview; 37% moved and could not be located. Differences between the initial sample and those retained for the second phase suggest some biases exist in the final sample. The individuals who responded to both questionnaires were dissimilar from the original sample in that they were more likely to be older, to have a higher income, to live in a single-family dwelling, and less likely to be black or to have children under six years of age. Additionally, fewer individuals were retained in the control sample (43%) than in the treatment group (57%). The effects of the sample mortality cannot be determined although it is likely that the individuals who were lost from the sample would be more responsive to ridesharing incentives than the older and wealthier travelers.

The individuals who were retained for analysis in the treatment group differed in several respects from the control group. The control group contained a significantly larger proportion of whites, hispanics, and sales workers than did the treatment group. Individuals from large households and dwellers in single-family residences were also overrepresented in the control group. The treatment group contained a larger proportion of blacks and clerical workers. The number and age of children, marital status and the number of licenses and vehicles in the household were not significantly different.

Results: Changes Over Time

Behavior

Individuals were assigned to a mode use category based on use of a mode for 60% or more of the work trips. The three categories of mode use were single-occupant auto (SOA), carpool (CP), and bus (BUS). Where an individual used two modes at about the same rate, he is defined as a dual mode user, an individual who uses three modes for a significant amount of traveling is designated as a mixed mode user. The proportion of trips made in each mode for the user groups at time 1 is presented in Tables 2 and 3. Over 90% of the treatment and control samples used one mode for most of their trips to work. The remainder divided their trips almost equally between two modes.

At time 1, 38% of the treatment group and 58% of the control group used one mode for 100% of their work trips. Almost a majority of the total sample used at least one other way of traveling during a monthly period. At time 2, a larger proportion of the sample had clearly defined an exclusive mode to work; 52% of the

treatment group and 61% of the control group used the same mode every day. A slight increase was also found in the number of people who alternated between two or more modes. In the treatment group, 7% initially used dual modes whereas at time 2, 10% combined modes; in the control group in the percentages are 4 and 6, respectively.

Table 2

Proportion of Trips Made by Mode User Groups in the Treatment Area at Time One

Modal Groups	N	Proportion of Trips by ...		
		SOA	CP	BUS
SOA	109	94.3%	4.4%	1.2%
CP	36	5.7%	92.7%	1.6%
BUS	52	4.9%	2.9%	92.2%
SOA-CP	9	47.6%	49.8%	2.6%
SOA-BUS	4	48.0%	3.9%	48.0%
CP-BUS	2	0.0%	50.0%	50.0%
SOA-CP-BUS	0	-	-	-
	<u>N=212</u>			

Table 3

Proportion of Trips Made by Mode User Groups in the Control Areas at Time One

Modal Groups	N	Proportion of Trips by ...		
		SOA	CP	BUS
SOA	114	97.0%	1.8%	1.2%
CP	65	4.0%	93.8%	2.2%
BUS	84	2.4%	1.3%	96.4%
SOA-CP	5	47.4%	51.4%	1.2%
SOA-BUS	4	50.0%	0.0%	50.0%
CP-BUS	2	0.0%	50.0%	50.0%
SOA-CP-BUS	0	-	-	-
	<u>N=279</u>			

Changes in the choice of mode to work can be observed in Tables 4 and 5. In the treatment group, 72% of the travelers used the same mode for a majority of their trips at both time periods. Most of the change in the treatment group is due to mode users combining their time one mode with another at time two, or the dual mode users at time one selecting one of those two modes for greater use at time two. Only 13% switched to a mode rarely used at time one. Switching from single-occupant autos was less pronounced than defections from carpools and buses; 19% of those who drove alone at time one switched to carpools, buses or combined modes while 36% of the carpools and 23% of the bus users altered their mode choice behaviors. In the control group, 80% of the respondents used the same mode at both points in time and 13% changed modes radically. Eleven percent of the time one single-occupant auto drivers altered their mode choice behavior compared with 31% of the carpools and 17% of the bus riders.

The small number of dual and mixed mode users necessitated their assignment to a modal group on the basis of the mode used most often in traveling to work. The average mode frequency was then analyzed to determine the significance of changes for time 1 mode-use groups (Tables 6 and 7). In the treatment group, single-occupant auto drivers significantly reduced their use of the mode at time 2 while increasing their use of both buses and carpools. Bus and carpool users significantly decreased the average frequency of use for their respective modes. Carpools more frequently drove alone and took buses at time 2; bus users increased their use of the carpool ($p = .056$). The same pattern

was observed in the control group; mode use groups from time 1 decreased their average use of their respective modes at time 2. Single-occupant auto, bus and carpool users increased their use of the two alternative modes.

Mode choice behavior is based on trial and evaluation of alternative modes for much of the sample; at time 1, 38% of the treatment group and 58% of the control group used a second mode at least once during a monthly period. The single-occupant auto drivers' experience with ridesharing modes was surprisingly high; 75% of the treatment and 73% of the control auto users had taken a bus or a carpool at least once a month. While mode choice behavior changed over time in both treatment and control groups, only 14% of the time 1 single-occupant auto users took a ridesharing mode for more than half of their trips at time 2. This fact combined with the significant level of defections from carpools and buses reaffirms what is already known: ridesharing is not a particularly satisfactory alternative to driving alone for the vast majority of the traveling public.

Contrary to expectations, the mode distribution at time 2 was approximately the same as at time 1. No significant net decrease in driving alone was observed in either the treatment or control groups as a whole although a significant reduction in the frequency of driving alone was obtained for the time 1 single-occupant auto users. This suggests that some travelers who drive alone can be convinced to rideshare. However, the significant defections from the time 1 ridesharing groups indicate that retaining the carpools and bus users to produce a net increase in ridesharing is problematic.

Perceptions

The modal attributes which changed between measurement points are presented in Table 8. The single-occupant automobile attributes which were perceived to change, all changed in the negative direction for both the treatment group and the control group. For the treatment group as a whole, bus travel time, crowding and waiting in traffic improved over time. Carpool attributes of crowding, package space and ease of use changed negatively. Carpool travel time and waiting in traffic significantly improved and were subsequently more positive than those attributes for the alternative modes. Bus safety and cost were perceived more negatively at time 2 in the control group, although waiting in traffic significantly improved over time for the carpool mode.

The changes in perceptions for the total treatment and control groups indicate that some attributes had changed for all mode use groups. In the treatment group, changes are those expected: the attributes of driving alone were worse at time 2, while bus and carpool attribute perceptions of travel time and waiting in traffic improved over time. In the control group, perceptions of the bus and single-occupant automobile were lower at time 2 than at time 1. When the sample was partitioned into subsets in terms of time 1 mode use, some changes in variables which were significant for the total sample were found to be insignificant within any one group (Tables 9 and 10). Furthermore, perceptions of other attributes were found to have significantly altered over time for certain user groups. Within the treatment group, only single-occupant automobile users perceived travel time in that mode to be significantly worse at time 2. No other significant differences were observed in single-occupant auto perceptions of the mode. Bus travel time, bus crowding, bus exposure to weather, and bus waiting in traffic significantly improved although they were still not perceived as strongly positive (4.1). The wait in traffic for the carpool was perceived more positively

Table 4

Changes in Mode Choice by Time One User

Groups in the Treatment Area

Time One Mode Choice

Time 2										
Mode Choice	SOA	CP	BUS	SOA-CP	SOA-BUS	BUS-CP	MIXED	% of t2	N	
SOA	80.7%	11.1	9.6	66.7	50.0	0.0	0.0	49.5	105	
CP	5.5%	63.9	5.8	22.2	0.0	50.0	0.0	16.5	35	
BUS	5.5%	5.6	76.9	0.0	0.0	50.0	0.0	24.1	51	
SOA-CP	2.8%	5.6	0.0	11.1	25.0	0.0	0.0	3.3	7	
SOA-BUS	3.7%	0.0	3.8	0.0	25.0	0.0	0.0	3.3	7	
CP-BUS	0.0%	8.3	3.8	0.0	0.0	0.0	0.0	2.4	5	
MIXED	1.8%	0.0	0.0	0.0	0.0	0.0	0.0	0.9	2	
% of t1	51%	17%	25%	4%	2%	1%	0.0%			
N	109	36	52	9	4	2	0		212	

Table 5

Changes in Mode Choice by Time One User

Groups in the Control Areas

Time One Mode Choice

Time 2										
Mode Choice	SOA	CP	BUS	SOA-CP	SOA-BUS	BUS-CP	MIXED	% of t2	N	
SOA	88.6	18.5	5.6	40.0	0.0	0.0	0.0	43.0	120	
CP	3.5	69.2	5.6	40.0	0.0	0.0	0.0	20.1	56	
BUS	4.4	7.7	83.1	0.0	50.0	50.0	0.0	31.2	87	
SOA-CP	0.9	1.5	0.0	20.0	25.0	0.0	0.0	1.4	4	
SOA-BUS	0.9	0.0	2.2	0.0	25.0	0.0	0.0	1.4	4	
CP-BUS	0.0	3.1	2.2	0.0	0.0	50.00	0.0	1.8	5	
MIXED	1.8	0.0	1.1	0.0	0.0	0.0	0.0	1.1	3	
% of t1	40.9	23.3	31.9	1.8	1.4	0.7	0.0			
N	114	65	89	5	4	2	0		279	

at time 2 than at time one. However, the carpool attributes of package space and crowding were perceived to have worsened by single-occupant automobile users in the treatment group. Time one carpoolers believed that at time 2, carpool travel time has significantly improved. Bus travel time and waiting in traffic improved but were still slightly negative (≤ 3.8) while extra time for the bus became marginally positive. Driving alone was considered to be more costly at time 2 than at time 1.

For treatment group bus users, the cost of bus use improved while the reliability of the bus decreased over time. Bus users believed the carpool was less

easy to use at time 2 but the wait in traffic and parking cost of the mode had improved. No significant changes were observed in the bus users' perceptions of driving alone.

Within the control group, single-occupant automobile users believed the ease of use, the ease to the destination, personal safety and waiting in traffic had significantly worsened for driving alone. Attributes for the carpool and bus modes also significantly decreased over time. The carpoolers believed that the carpool mode had increased in reliability; their perceptions of bus safety and cost were more negative at time 2. Control group bus riders believed the bus was

less comfortable at time 2 and more costly than at time 1 and that carpool convenience had decreased between the pre-test and post-test.

Table 6

Change in Proportional Use of Modes Time 1-Time 2: Treatment Group

Mode Used at Time 1		Proportion of Trips by...		
		SOA	CP	BUS
SOA Users	time 1	91.2%	5.9%	2.9%
	time 2	80.9%	10.4%	8.7%
	(p)	(.001)	(.036)	(.009)
CP Users	time 1	10.2%	86.0%	3.9%
	time 2	22.8%	61.9%	15.3%
	(p)	(.006)	(.000)	(.017)
BUS Users	time 1	4.9%	2.9%	92.2%
	time 2	12.4%	8.7%	78.8%
	(p)	n.s.	(.056)	(.006)
All Users	time 1	53.6%	21.4%	25.0%
	time 2	52.3%	20.4%	27.3%
	(p)	n.s.	n.s.	n.s.

Table 7

Change in Proportional Use of Modes Time 1-Time 2: Control Group

Mode Used at Time 1		Proportion of Trips by...		
		SOA	CP	BUS
SOA Users	time 1	94.3%	3.0%	2.8%
	time 2	84.5%	7.1%	7.6%
	(p)	(.000)	(.020)	(.014)
CP Users	time 1	5.0%	91.4%	3.6%
	time 2	21.4%	66.6%	12.0%
	(p)	(.000)	(.000)	(.004)
BUS Users	time 1	2.4%	1.3%	96.4%
	time 2	8.1%	7.9%	84.0%
	(p)	(.012)	(.008)	(.000)
All Users	time 1	42.9%	24.3%	32.8%
	time 2	44.5%	22.1%	33.1%
	(p)	n.s.	n.s.	n.s.

Table 8

Significant Attribute Changes from Time 1 to Time 2 (All Mode Users)

Control Group		t1 mean	t2 mean
Mode Attributes			
Carpool	Wait in Traffic	3.6	3.9
	Ease of Use	6.4	6.2
	Personal Safety	5.6	5.4
BUS	Cost	5.7	5.2
	Vehicle Safety	5.7	5.5
	Personal Safety	5.0	4.6
Treatment Group			
Carpool	Package Space	5.0	4.6
	Ease of Use	4.3	3.8
	Crowding	4.6	4.3
	Travel Time	4.1	4.7
SOA	Wait in Traffic	3.6	4.4
	Travel Time	4.3	3.7
	Wait in Traffic	3.5	3.1
BUS	Flexible Schedule	6.5	6.2
	Cost	5.8	5.5

Table 8 continued

Mode Attributes	t1 mean	t2 mean
Travel Time	2.7	3.6
Crowding	2.8	3.3
Wait in Traffic	3.3	4.1

Table 9

Pre-test to Post-test Perception Changes Treatment Group

Time 1 Mode	Perception	Mean t1	Mean t2
SOA Users	Carpool Package Space	5.0	4.5
	Carpool Crowding	4.5	4.0
	Carpool Wait in Traffic	3.7	4.5
SOA Users	SOA Travel Time	4.6	3.9
	Bus Travel Time	2.6	3.4
	Bus Crowding	2.4	3.1
	Bus Weather	2.7	3.2
	Bus Wait in Traffic	3.2	4.1
Carpool Users	Carpool Travel Time	3.8	5.1
	SOA Parking Cost	4.2	3.3
Bus Users	Bus Travel Time	2.2	3.4
	Bus Wait in Traffic	2.8	3.8
	Bus Extra Time	3.2	4.1
	Carpool Ease of Use	4.8	3.7
	Carpool Wait in Traffic	3.3	4.2
	Carpool Parking Cost	4.7	5.7
	Bus Reliability	5.7	5.0
Bus Parking Cost	6.2	6.8	

Changes Significant at (p<.05).

Table 10

Pre-test to Post-test Perception Changes Control Group

Time 1 Mode	Perception	Mean t1	Mean t2
SOA Users	Carpool Package Space	4.7	4.2
	Carpool Ease to Destination	5.2	4.5
SOA Users	Carpool Crowding	4.9	4.2
	Carpool Vehicle Wait	4.7	4.1
	SOA Ease of Use	6.8	6.5
	SOA Personal Safety	6.1	5.5
	SOA Ease to Destination	6.9	5.6
	SOA Wait in Traffic	4.5	4.1
Bus Users	Bus Cost	5.4	4.8
	Bus Vehicle Safety	5.6	5.2
	Bus Crowding	2.6	2.1
Carpool Users	Carpool Reliability	5.1	5.8
	Bus Cost	5.5	4.7
Bus Users	Bus Personal Safety	5.3	4.5
	Carpool Convenience	4.5	3.8
	Bus Comfort	5.5	5.1
	Bus Cost	6.4	5.9

Changes significant at (p<.05).

For both the treatment and control groups, changes in single-occupant automobile and bus group perceptions corresponded with behavior. The decrease in use of the carpool by time 1 carpoolers was not indicated by the aggregate perceptions.

Changes in Attribute Importance

Perceptions of mode attributes may change and thereby induce behavior change; it is also possible that the

salience of attributes changes over time and affects behavior. The value that individuals use in determining mode choice may be altered by changes in circumstances or reference groups, because basic motives have changed, because attribute importance is rationalized to correspond with behavior or because an attribute which was "taken for granted" or outside the range of possibility, is altered or becomes available. To evaluate changes in the importance of an attribute in determining behavior, the unstandardized regression coefficients from simple regressions were analyzed (Table 11). Simple, rather than multiple, regressions were used because of known multicollinearity in the data and entry into the multiple regression depends on residual unexplained variance rather than total unexplained variance. A t test was applied to test for significant differences over time.

Within the treatment group, the attributes of arriving on time, vehicle safety and extra time became more important over time in the determination to take a carpool. The increased salience of safety and extra time could be due to the high accident rate on the Santa Monica Freeway and the requirement of three vehicle occupants to use the reserved lane. The increased importance of arriving on time in the decision to use either a single-occupant auto or a carpool may also be due to negative impacts of the reserved lane. Where arriving on time at time 1 had been a matter of personal responsibility in both automotive modes, at time 2 the unpredictable nature of the congestion on the freeway determined whether one arrived on time.

Ease of use in driving alone changed significantly in importance. The media representation of carpools and buses may have served as a negative referent by making that aspect of driving alone even more advantageous or it may be due to unexpected difficulties in use of the mode at time 2. Safety factors were more influential in the determination to take a bus at time 2 than at time 1. Again, the increase in traffic accidents may have highlighted this attribute. Alterations in the importance of parking cost also occurred; the fare increase may have focused individual attention on the lack of a parking cost as a way of rationalizing the cost savings usually attributable to the bus mode. Extra time was found to become more important while crowding decreased in importance. These changes may have resulted from the improvements in bus service which contrasted with what had been available.

In the control group, vehicle safety became more important and personal safety became less important in the determination to drive alone. The reported accident rates on the Santa Monica Freeway may have highlighted an attribute which had little meaning at time 1 and replaced the more traditional perception of the auto as a guarantor of personal safety. Changes in the importance of waiting in traffic, relaxing and exposure to weather in the decision to take a bus may have resulted from the improved bus service on the freeways. The causal structure with respect to the carpool mode changed significantly in the control group. Eight of the 19 attributes became more important over time.

Media coverage of events which occurred between measurements may have aroused interest in the mode and provided information where attitudes were not strongly held. The interview at time 1 may also have stimulated reassessment of the shared ride. Furthermore, trial of the mode would inform individuals about the characteristics of the mode, and if discrepant with expectations would generate reevaluation of the carpool and the value of its attributes.

Importance of the attributes can change through alterations in the external environment and through changes

in individual values, perceptions or behaviors. Individuals may be indifferent to certain attributes at time 1 because they were associated with all 3 modes or were considered necessary or neutral qualities of a particular mode. Where the modes were later differentiated with respect to the attribute or where the attribute was found to significantly change, the individual may reassess its importance. Comparisons between the modes engendered through communications or personal experiences would serve to emphasize the attribute levels associated with usage of a particular mode. The changed reference structure may induce reevaluation of the importance of the characteristics of the modes. Furthermore, changing the importance of attributes may also represent dissonance reduction where behavior and perceptions were not congruent at time 1.

The Effect of the Treatment on Perceptions and Behavior

In order to assess the effects of the reserved lane on mode user groups, subjects in both the treatment and control groups were divided into three categories, on the basis of which mode-to-work (single-occupant auto, carpool, or bus) was most frequently used prior to the modifications in the Santa Monica Corridor. Changes in beliefs and behavior were then analyzed separately for the three groups.

A significant difference between the beliefs in the treatment group vis-a-vis the control group was observed in 15 of the 157 comparisons. These results are summarized in Table 12.

For ten attributes, differences between the control and treatment groups involve significant changes over time in the perceptions of at least one of the groups. The remaining five attributes, while not significantly different from time 1 to time 2 within the groups, display a shift in opposing directions for the treatment and control groups. The difference in the pattern of change suggests some treatment effect. The treatment would be expected to affect bus and carpool perceptions by making them more positive and to reduce the positive perceptions of driving alone. Ten of the 15 perceptual changes are in the expected direction for the treatment group.

For the time 1 single-occupant auto users, the treatment produced a significant change in the perception of the wait in traffic required for carpools. The attribute changed from slightly negative to positive for the treatment group while remaining negative in the control group. For the treatment group, carpool waiting in traffic was more positive at time 2 than perceptions of the wait in traffic necessitated by driving alone or riding in a bus. Most of the remaining perceptions of the attributes of the carpool were less positive than the comparable beliefs about single-occupant autos. No other modifications of carpooling beliefs were significantly different in the treatment group in comparison with the control group.

The single-occupant automobile treatment group perceived that the travel time of the mode had significantly increased from time 1 to time 2, while the control group did not alter their perceptions of this attribute. Both the perceptions of the carpool wait in traffic and the increased travel time incurred in driving alone can be attributed to the changes on the Santa Monica Freeway. However, a significant difference between the treatment and control groups with respect to single-occupant automobile personal safety and ease to destination was unexpected. There was no significant change in the treatment group but the control group perceptions of personal safety and ease to the destination decreased significantly. The actual cause of the change in perceptions for the control group is

Table 11
 Changes in Importance of Attributes in Explaining Behavior
 Unstandardized Regression Coefficients

Attribute	Treatment Group		Control Group		
	t1	t2	t1	t2	
CP	Comfort	.618	.889	.419	1.09
	Package Space	.490	.615	.497	1.10
	Ease of Use	.739	1.174	.796	1.35
	Reliability	.649	1.135	.618	1.58
	Arrive on Time	.405	1.334	.783	1.50
	Vehicle Safety	.294	.948	.992	.982
	Crowding	.846	.699	.213	.797
	Waiting Time	.634	1.168	.369	1.34
	Weather	.593	.481	.466	1.05
	Extra Time	.673	1.315	.886	.768
SOA	Ease of Use	.881	1.66	1.41	1.13
	Arrive on Time	-.209	.845	1.24	.990
	Vehicle Safety	1.23	.809	.518	1.06
	Personal Safety	.513	.075	1.10	.521
BUS	Vehicle Safety	.421	1.23	.844	.943
	Personal Safety	-.134	.578	.557	.561
	Crowding	1.48	.760	.837	1.02
	Waiting Time	1.27	1.50	1.66	2.40
	Relaxation	.856	1.24	1.21	1.75
	Weather	.983	.543	.750	1.25
	Extra Time	.833	1.34	1.71	1.79
	Parking Cost	-.953	1.25	-.130	-.782

All differences significant at $p \leq .05$.

Table 12
 Treatment Effects Upon Perceptions
 of Attributes

Time Mode 1	Mode Perceptions	Treatment $\bar{X}_{t1} - \bar{X}_{t2}$	Control $\bar{X}_{t1} - \bar{X}_{t2}$
SOA Users	CP Waiting in Traffic	3.7-4.5*	3.6-3.8
	SOA Travel Time	4.6-3.9*	4.4-4.5
	SOA Personal Safety	5.8-5.7	6.1-5.5*
	SOA Ease to Destination	6.6-6.7	6.9-5.6*
	Bus Crowding	2.4-3.1*	2.6-2.1*
	Bus Waiting for Vehicle	3.2-3.5	3.3-3.0
	Bus Exposure to Weather	2.7-3.2*	2.8-2.6
CP	CP Easy to Use	5.2-4.5	5.1-5.4
	CP Crowding	5.4-5.1	5.1-5.6
	CP Travel Time	3.8-5.1*	4.9-5.1
	SOA Wait in Traffic	3.3-2.6	3.8-4.4
	Bus Travel Time	2.2-3.4*	3.3-3.1
BUS	CP Easy to Use	4.8-3.7*	4.5-4.4
	SOA Extra Time	5.9-6.4	6.3-6.0
	Bus Parking Cost	6.2-6.8*	6.4-6.0

Items coded (1) least positive to (7) most positive

Treatment effects significant at $p \leq .05$ level (two-tailed)

(*) Starred within group differences significant at $p \leq .05$ level (two-tailed)

unclear and illustrates one problem with the non-equivalent control group design. Any event which affected those living in the control areas but not the treatment zone could result in a confounding of the treatment effect. The treatment group did reduce its frequency of single auto use significantly; however, the control group also reduced its single-occupant auto use by about the same amount. Thus, it is unclear whether the drop

in the treatment group was caused by the policy change and the decrease in the control group by some other factor; or whether both decreases are a product of some other cause.

For single-occupant automobile users, three bus beliefs were found to differ significantly between the treatment and control groups. Bus crowding, waiting for the bus

and exposure to weather were significantly more positive at time 2 than at time 1 for the treatment group. Transportation system modifications which decreased bus headways and increased frequency of bus service in the treatment zones would, therefore, seem to have had the desired effect upon this group. However, changes in the behavior of the treatment group were not significantly different from changes in the control group. Bus use increased for time 1 single-occupant auto users in both samples.

There was a significant treatment effect on the perceptions of the carpool users; the direction of change was, however, contrary to expectation. The treatment group showed a decrease in the ease of use of the carpool mode while the control group found the mode to be improving in its ease of use. This may be a reflection of the reported difficulty in crossing traffic lanes to get to the diamond lane or confusion over the new policy. Perception of crowding in carpools also was less favorable in the treatment group at time 2 and may be due to the need to have at least 3 people in the carpool in order to use the reserved lane. Travel time during rush hour in carpools moved in the desired direction; the treatment group exhibited an increasingly positive perception for this attribute.

The only significant treatment effect on the carpoolers' perceptions of driving alone was a relative increase in the time spent waiting in traffic for the treatment group, while the control group perceived a decrease in waiting time for the single-occupant auto. This is probably attributable to the Santa Monica ramp metering system and the use of ramp bypasses by the carpoolers.

The carpoolers in the treatment group also perceived a significant improvement in bus travel time while the control group perceptions were slightly worse at time 2 with respect to the attribute of the bus. The mean bus use of the carpoolers increased equally in each group.

Bus users in the treatment group considered the carpool to be less easy to use at time 2 than at time 1, and, thereby differed significantly from the control group. Again, this may be attributable to the observed difficulties in accessing the reserved lane. A treatment effect was also observed in the extra time required in driving alone. The treatment group changed positively with respect to this attribute while the control group changed to a less positive perception of single-occupant automobile extra time. Some confusion may have existed with respect to the definition of the attribute; individuals may have changed their parking locations or may be including the extra time required after leaving their car at a park-and-ride lot.

A treatment effect was observed for bus parking cost among the bus users. The more positive belief observed at the post-test for the treatment group is perhaps a reflection of reduced costs for those using a park-and-ride lot. Finally, no significant effect was observed in the behavior of bus users.

No significant treatment effect upon behavior was observed. While significant changes were found in the treatment zone in use of all three modes, behavior change was occurring at the same level in the control group. Several explanations can be offered. First, it is likely that exposure to media and to acquaintances who used the Santa Monica Freeway reserved lane heightened interest in the ridesharing modes and the transportation system in general for the control group. Individuals who were potential ridesharers may have become more aware of their alternatives and availed themselves of the opportunity to try a carpool or bus. Secondly, the definition of treatment as residence in the Santa Monica corridor may be too broad. It is

likely that the reserved lane was perceived as an incentive only for those who could use the freeway for the entire 12.5 miles. Third, the reserved lane on the Santa Monica Freeway was controversial, and consistent threats to its continuation undoubtedly mitigated against the stabilization of attitudes and behaviors.

The changes that were observed in the panel of downtown commuters thus appear to be attributable to a dynamic ridesharing segment of the population for whom increased awareness and availability of the ridesharing mode facilitated the use of the alternatives to driving alone to work.

Conclusions

A significant change in mode choice behavior was observed in both the treatment and control groups over time. In each group, however, only 13% changed to modes with which they had had little experience at time one. In the treatment group, single-occupant auto, carpool and bus users at time 1 decreased their use of their respective modes at time 2. Single-occupant auto users significantly increased their use of the bus and carpool while carpoolers increased their use of the bus and single-occupant auto. Bus users significantly increased their use of the carpool. The same pattern is found in the control group; time 1 mode groups decreased their use of those modes at time 2. Both carpoolers and single-occupant auto users increased their use of the bus at time 2; carpoolers and bus users also increased the use of the single-occupant auto. Both time 1 bus and single-occupant auto users increased their carpool usage.

It is likely that in the process of choosing a mode, individuals try at least one of the alternatives. A considerable number of individuals used two modes about equally in traveling to work. However, those who alternated between two modes at time one had generally selected between them by time two. Furthermore, a majority of the sample used a second way of traveling to work at least once at time one. It appears that most of the Los Angeles downtown commuters are not captive to a mode and that most have experienced at least one of the ridesharing alternatives. The significant defections from the bus and carpool modes suggest a dynamic ridesharing population. Continual promotional and information campaigns are necessary to continue to attract commuters and to thereby replace those who have switched.

Contrary to expectations, the treatment group did not display a large increase in the use of the carpooling mode. Travel counts on the Santa Monica Freeway revealed a 65% increase in carpools and a 300% increase in bus use (Billheimer, et al., 1977). This travel data is derived with respect to all the riders of the freeway, while the survey sample included only downtown workers. It is possible that downtown commuters were already ridesharing in greater proportions than the non-CBD workers or that many who had relocated had significantly altered their travel behavior. It is also likely that the survey sample was not representative of the total population. In the random survey of households the number of downtown workers and their mode to work were obtained; 56% of the total household workers drove alone to work, 20% used a carpool and 24% were bus users. This mode distribution is approximately equal to the mode distribution of the time 1 treatment group that was retained for analysis. It is possible, however, that the background characteristics of the mode users in the sample are significantly different from the population and, therefore, explain the lack of change observed.

The reserved lane on the Santa Monica Freeway appears to have had relatively little effect on the behavior of the downtown commuter sample. Behaviors in the control group were not significantly different than those in the treatment group. However, one of the problems in the analysis is the inability to separate the increased use of the bus associated with express buses using a reserved lane from the mere availability of express buses. Increased bus service in all areas combined with additional express routes on control freeways confounded the analysis. The only conclusion which is indicated is that the reserved lane implemented on the Santa Monica Freeway for five months did not increase carpooling behavior among the downtown commuters. Whether this lack of carpooling behavior can be attributed to the controversial nature of the reserved lane and the continuous discussion of its termination, cannot be determined with the data available.

The importance of modal attributes, as measured by their influence upon behavior, changed over time in both the treatment and the control groups. System modifications, communication, and personal experience with the modes are presumed to have caused the reassessment of attribute salience. Reactions to alterations in service as well as the enhanced comparative evaluation of the attribute levels between modes contributed to the change in the importance of attributes in the mode choice decision.

Perceptions of the modes were also altered in both the treatment and control groups and in general corresponded with the behavior patterns. Significant differences in perceptions were observed in the treatment and control groups. Single-occupant automobile treatment perceptions were found to be more positive than single-occupant automobile control group perceptions of the bus and carpool modes and although the attributes of carpool and bus remained generally less positive than those of the single-occupant auto mode, the magnitude of difference was decreased. It is possible that had all the incentives to multi-passenger vehicle use remained in effect, further behavior changes in the direction of a more balanced mode split might have been attained.

The relationship between behaviors and perceptions of modal attributes can only be determined with individual-level analysis. Further research is required regarding the effects of perceptual consistency and behavior change, changed availability and mode choice, communication flows and mode trial, as well as the determination of the modal attributes which appear to influence behavior.

References

- John W. Billheimer, Robert Bullemer and Carolyn Fratessa, (SYSTAN, Inc.) "The Santa Monica Diamond Lanes: An Evaluation," preliminary report to the U.S. Department of Transportation. Transportation Systems Center, Contract No. DOT-TSC-1084, (Washington, D.C.: 1977).
- James L. Bruning and B. L. Kintz, Computational Handbook of Statistics, (Glenview, IL: Scott, Foresman, 1968) p. 13-15.
- Donald T. Campbell and Julian C. Stanley, Experimental and Quasi-Experimental Designs for Research, (Chicago: Rand McNally, 1963).
- Gorman Gilbert and James F. Foerster, "Public Response to the Chapel Hill Bus System," Department of City and Regional Planning, University of Carolina (Chapel Hill: 1975).
- William Goss and Paul Shuldiner, "University of Massachusetts Bus Research and Demonstration Grant" awarded 1972 Mass-DMG-6 by UMTA Research and Development, Final Report expected 1977.
- David T. Hartgen and Stephen M. Howe, "Irondequoit-Wayne Expressway: Before-After Study, Design Report," Preliminary Research Report 101 (Albany, New York: Planning Research Unit, New York Department of Transportation, June 1976).
- David A. Kenny, "A Quasi-experimental Approach to Assessing Treatment Effects in the Non-Equivalent Control Group Design," Psychological Bulletin, 82:3, (May 1975) p. 345-362.
- Robert L. Linn and Charles E. Werts, "Analysis Implications of the Choice of A Structural Model in the Non-equivalent Control Group Design," Psychological Bulletin, 84:2 (March 1977) p. 229-234.
- Daniel McFadden, Antti Talvitie, Stephen Cosslett, Ibrahim Hasan, Michael Johnson, Fred A. Reid, and Kenneth Train, "Demand Model Estimation and Validation," Institute of Transportation Studies, University of California, Berkeley, California (June 1977).
- Quinn McNemar, Psychological Statistics, Fourth Edition, (New York: John Wiley and Sons, 1969), p. 158.
- Thomas Parody, "An Analysis of Disaggregate Mode Choice Models in Prediction," paper presented at annual meeting of Transportation Research Board, (1977).
- Louise Skinner, "The Effect of Energy Restraints on Travel Patterns," Urban Planning Division, Federal Highway Administration (Washington, D.C.: 1975).
- Peter L. Watson and Edward P. Holland, "Road Pricing in Singapore: Impacts of the Area License Scheme," draft report, Urban Projects Department, World Bank (1977).

HALO EFFECTS IN MARKETING RESEARCH: REVIEW AND PROGNOSIS

Neil E. Beckwith, University of Pennsylvania
Harold H. Kassarian, University of California, Los Angeles
Donald R. Lehmann, Columbia University

Abstract

The tendency in rating an object on a particular attribute to be influenced by a general impression is well known. Some of the early psychological studies of this halo effect are reviewed along with more recent work in marketing and consumer behavior.

Introduction

Consumers' ratings of a brand (or other object) on a set of dimensions (attributes or traits) are commonly used in consumer research. Examples include image (i.e., profile) studies, advertising effects studies, and product attribute positioning studies. However, an individual's ratings on an attribute may be determined by many other variables besides cues directly relating to the particular attribute.

In a recent Federal Trade Commission case, consumers' perceptions of a brand of bread were at issue. A complaint of the FTC had alleged that ITT Continental Baking Company had falsely represented Wonder Bread to be superior in nutrition to competing brands of white bread.

It was agreed on both sides that some percentage of the public believed that Wonder Bread stood out in nutrition. One issue was the extent to which specific advertising about the nutritional quality of the brand had influenced these beliefs. The bakery company claimed that the nutritional beliefs about Wonder Bread were influenced by consumers' generally favorable overall attitudes toward the brand (including familiarity, price, and higher levels of advertising) and that this positive overall feeling about the brand in turn influenced the belief that Wonder Bread was nutritionally superior. (Eventually this particular charge was dismissed on other grounds.)

The tendency in rating an object on a particular attribute to be influenced by a general impression is widely recognized.

Concept

Rating responses of an object on a set of attribute dimensions is generally determined as a function of several related aspects. First, it is generally held that some adjustment process exists whereby overall attitude and beliefs about the object on the attributes influence each other, and the individuals' beliefs become adjusted as a result of the overall attitude and other beliefs. This process is neither good nor bad; it exists and must be accounted for.

Second, rating response about the object is influenced not only by true beliefs but also by a variety of measurement effects. This means that a respondent, when rating an object on an attribute, may be responding to stimuli rather than his true belief.

Definitions

Interestingly, within the research community substantial

disagreement exists concerning the definition, importance, circumstances, and even the existence of a halo effect.

Tiffin and McCormick (1965) consider the halo effect to be the domination of all other traits by one particular trait. Krech, Crutchfield, and Bellachey (1962), in a similar vein, point out that if a person is well liked, he will be rated as being very high or positive on all other specific traits. If a person is disliked, perceptions of even his most positive attributes will be downgraded. The perception of the person is influenced by the overall feeling toward that person.

English (1934) many years ago provided a rather general definition of the halo effect as the "tendency in rating to be influenced by general impression or attitude when trying to judge separate traits." This view is broader than that of Krech et al, or Tiffin and McCormick in that haloing is seen as an individual phenomenon and the source of bias is the individual's general (overall) impression or attitude, and includes the rating of any objects.

From a more technical point of view, it has also been common to view the halo effect as the excessive partial correlation between belief ratings (e.g., Symonds, 1925). However, some "natural" correlation between beliefs should be expected even without any halo effect. It is difficult to partition the correlation between the halo and natural components (Bingham, 1939). The Federal Trade Commission adopted a definition in its final opinion in the Wonder Bread case which included three sources of the halo effect: (1) widespread usage, (2) generalized advertising claims, and (3) claims stressing a different attribute (Docket 8860, p. 13).

Related Theories

Various cognitive balance theories (individuals' attempts to maintain beliefs which are consistent with each other) also suggest the existence of halo effects. Heider (1958) points out that if an individual likes both person A and object X, he would expect that his friend, person A, would also like object X. If not, the individual would experience cognitive imbalance—an uncomfortable state. "How can my friend whom I like dislike object X or politician Y or a third person. Either object X is not as good as I thought it to be, or my friend is crazy and not quite as likeable as I thought him to be" (see, e.g., Abelson et al, 1969, Feather, 1964; and Rosenberg et al, 1960).

Festinger (1964) calls this imbalance dissonance, again an uncomfortable state that must be corrected. In this unconscious "correction" process the fact that feelings toward one attribute or person or product can systematically influence another object or attribute is very similar to halo effect from some points of view. Nice people have nice attributes, and less nice people have less nice attributes (see, e.g., Venkatesan, 1973; and Wickland and Brehm, 1976).

As can be seen we have already confused the halo effect defined as "the overall impression influencing specific attributes" with popularity, familiarity, etc. Hence it may be worthwhile to examine what we wish to call halo-like effects.

Halo-Like Effects

A number of other sources of rating bias are similar in spirit to the halo effect, but are actually distinct manifestations.

Other Attribute Effects

Belief about an attribute could possibly influence an individual's rating on another attribute beyond its effect via the overall attitude component. For example, if gasoline mileage was totally irrelevant to some individual's overall attitude, it might still influence his rating of miles per fill-up. Many studies have confirmed that the physical attractiveness of a person influences other persons' ratings of that individual on many attributes, as well as influencing the overall attitude (e.g., Miller, 1970).

Peer Attitudes

An individual's belief response may also be biased by his understanding of his own local peers' beliefs about the object on the attribute or overall attitude. Such peer group conformity influences have been considered by Bourne (1956), Sherif and Sherif (1964), and Siegel and Siegel (1957) among others.

Popularity

Individuals may also be biased by their understanding of the general public's beliefs or overall attitudes about the object, as well as their own local peer groups. Individuals may favorably bias their ratings for well known brands and unfavorably bias ratings for less popular products, for example.

Familiarity

Similarly, individuals may systematically influence their ratings differently for more familiar objects than for less familiar objects. The direction (sign) of such influence is difficult to prejudice, since it would seem to depend upon the general satisfaction level realized during familiarization with the object as compared to the individual's expectations of unfamiliar objects. Koltuv (1962) concluded that the magnitude of halo effects decreases with increasing familiarity; although, James and Carter (1977) did not find a similar result in their recent analysis of 14 students' location preferences.

Cross Sectional Estimation

One of the difficulties in assessing the halo effect is estimating the magnitude of the effect. The most critical problem is estimating the individual's true beliefs.

Cross Sectional Bias

"Halo effect" generally denotes the bias of each individual's belief response. However, measurement problems also exist. Huber and James (1976) pointed out that simply averaging the belief responses across individuals (as in Beckwith and Lehmann, 1975) does not provide unbiased estimates of the underlying locations or values. This is because the objects or alternatives will usually have different fractions of the respondents favoring them. The average of belief responses for the popular objects should be more favorably biased than for less popular objects since more individuals will favorably bias the rating for popular objects than they will for unpopular objects. Researchers have attempted to attenuate this source of bias by separately comparing users and nonusers across brands, for example. Beckwith and Kubilius (1977) have attempted to estimate the true locations of objects to remove this

bias. Also see Bemmar and Huber (1977) for a discussion of the necessity for simultaneous estimation of belief and attitude components.

Correlates of the Level of Halo

Thorndike (1920) originally concluded that the magnitude of halo effects seems to be surprisingly large. Subsequent investigators have attempted to determine which circumstances seem to involve only small or negligible halo effects, and which seem to encourage larger halo effects. There is substantial disagreement among published studies concerning even which circumstances encourage haloing. Thus, the following should be considered to be candidate, rather than proven, circumstances which tend to evidence more haloing:

1. Low familiarity with the objects (Koltuv, 1962).
2. Ambiguous or subjective attributes (James and Carter, 1977; Beckwith and Lehmann, 1975).
3. High perceived popularity or usage of the objects.

In addition, many other circumstances exist which might be expected to influence the degree of haloing. These include:

1. The importance of an attribute. An important attribute might be expected to show evidence of a larger halo effect since the attribute would more strongly influence overall attitude. However, Beckwith and Lehmann (1976) found a slight (but not significant) tendency for halo effects to decrease with increasing attribute importance.
2. Questionnaire wording can affect the degree of haloing (Wilkie, McCann, and Reibstein, 1973).
3. Relative importance of the object or product class.
4. Personal characteristics such as relative intelligence, education, gestalt-proneness, or interest of the respondent in the product class.
5. Promotional strategies employed. Whether the advertising environment primarily uses very general or very attribute-specific communications and claims may influence halo.

Seemingly divergent results have been reported, even for the same product class. For example, the Roper Organization reported that "The lower your gas mileage, the more uncomfortable the seats turned out to be" (Marketing News, 1975). However, Moore and James (1977) found only a very small amount of haloing in 34 students' ratings of automobiles on several attributes, including gas mileage and comfort. We conclude that unraveling the halo-inducing circumstances will not be an easy task, particularly as the focus of inquiry moves to attitude and belief changes.

Summary

Substantial disagreement exists concerning just when and how much haloing occurs. However it seems clear that a serious degree of haloing occurs in at least some circumstances. Until the nature of these circumstances is better understood it is reasonable for researchers to accommodate the possibility of halo effects within any study using belief ratings on attributes. Since the degree of haloing is expected to vary between individuals, individual level analysis would seem to be appropriate, at least until the halo effect phenomena is better understood.

Lastly, it is worth noting that in recent experimental manipulations, subjects were found to halo extensively, even when they had sufficient information to allow for independent assessments on the attributes. Furthermore, the subjects actually believed that the influence ran the other way, i.e., that their assessments on the attributes prompted their overall evaluative rating (Nesbill and Wilson, 1977; Ryan, 1977).

References

- R. P. Abelson, E. Aronson, W. J. McGuire, T. N. Newcomb, and M. J. Rosenberg, eds., Theories of Cognitive Consistency: A Sourcebook (Skokie: Rand McNally, 1969).
- Neil E. Beckwith and U. Victor Kubilius, "Empirical Evidence of Halo Effects in Store Image Research," Advances in Consumer Research (1977 ACR Proceedings).
- Neil E. Beckwith and Donald R. Lehmann, "The Importance of Halo Effects in Multi-Attribute Attitude Models," Journal of Marketing Research, 12 (August 1975), 265-275.
- _____, "Halo Effects in Multiattribute Attitude Models: An Appraisal of Some Unresolved Issues," Journal of Marketing Research, 13 (November 1976), 418-421.
- Albert C. Bemmar and Joel C. Huber, "Econometric Estimation of Halo Effects: Single vs. Simultaneous Equation Models," Advances in Consumer Behavior (1977 ACR Proceedings).
- W. V. Bingham, "Halo, Invalid and Valid," Journal of Applied Psychology, 23 (1939), 221-229.
- F. S. Bourne, "Group Influence in Marketing and Public Relations," in R. Likert and S. P. Hayes, eds., Some Applications of Behavioral Research, UNESCO, Paris, 1957, 207-257.
- Horace B. English, A Student's Dictionary of Psychological Terms (New York: Harper Brothers, Inc., 1934, fourth edition).
- N. T. Feather, "A Structural Balance Model of Communication Effects," Psychological Bulletin, 71 (1964), 291-313.
- Leon Festinger, Conflict, Decision, and Dissonance (Stanford: Stanford University Press, 1964).
- Fritz Heider, The Psychology of Interpersonal Relations (New York: John Wiley & Sons, 1958).
- Joel C. Huber and William L. James, "A Theory of Halo," working paper, Krannert School of Industrial Administration, Purdue University, March 1976.
- _____, "A Concept of Halo," Proceedings of the Association for Consumer Research, 1977.
- William L. James and Forrest G. Carter, "Halo Effects and Location Preferences," Advances in Consumer Research (1977 ACR Proceedings).
- Barbara Koltuv, "Some Characteristics of Intrajudge Trait Intercorrelations," Psychological Monographs, 76 (1962).
- David Krech, Richard S. Crutchfield, and Egerton L. Ballachey, Individual in Society: A Textbook of Social Psychology (New York: McGraw Hill, 1962), 67.
- Marketing News, "Marketing Briefs," June 6, 1975, 5.
- Arthur G. Miller, "Role of Physical Attractiveness in Impression Formation," Psychonomic Science, 19 (1970), 241-243.
- William L. Moore and William L. James, "A Study of the Amount of Halo in the Perceptions of Automobiles," Advances in Consumer Behavior (1977 ACR Proceedings).
- Richard E. Nesbitt and Timothy DeCamp Wilson, "The Halo Effect: Evidence for Unconscious Alteration of Judgments," Journal of Personality and Social Psychology, 35 (1977), 250-256.
- M. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson, and J. W. Brehm, Attitude Organization and Change, An Analysis of Consistency Among Attitude Components (New Haven: Yale University, 1960).
- Michael J. Ryan, "An Experimental Investigation into the Effects of Halo on Attitudinal and Normative Structures," Advances in Consumer Behavior (1977 ACR Proceedings).
- M. Sherif and C. W. Sherif, Reference Groups, Exploration into Conformity and Deviation of Adolescents (New York: Harper and Row, 1964), 170-172.
- A. E. Siegel and S. Siegel, "Reference Groups, Membership Groups, and Attitude Change," The Journal of Abnormal and Social Psychology, 55 (1957), 360-364.
- Percival Symonds, "Notes on Ratings," Journal of Applied Psychology, 9 (1925), 188-195.
- E. L. Thorndike, "A Consistent Error in Psychological Ratings," Journal of Applied Psychology, 4 (January 1920), 25-29.
- Joseph Tiffin and Ernest J. McCormick, Industrial Psychology (Englewood Cliffs: Prentice Hall, 1965), fifth edition.
- M. Venkatesan, "Cognitive Consistency and Novelty Seeking," in Scott Ward and Thomas S. Robertson, eds., Consumer Behavior: Theoretical Sources (Englewood Cliffs: Prentice Hall, 1973), 354-384.
- Robert A. Wickland and Jack W. Brehm, eds., Perspectives in Cognitive Dissonance (New York: L. Erlbaum Assoc., 1976).
- William L. Wilkie, John M. McCann, and David L. Reibstein, "Halo Effects in Brand Belief Measurement: Implications for Attitude Model Development," Advances in Consumer Research, I (1973 ACR Proceedings), 280-290.

A MEASURE OF HALO

Joel Huber, Columbia University
William James, University of Alabama

Abstract

Several different ways to measure halo are reviewed and a new measure proposed which is logically consistent and simple to implement. The use of familiarity judgments as an intervening variable is proposed as a way to estimate the degree of directional flow in halo judgments.

Introduction

Halo effect appears to be one of those concepts which (like a rainbow) is clear at a distance but dissolves on closer inspection. Halo occurs where affect or preference for an object biases attribute judgments in such a way as to make the judgments more consistent with preference. It may be lamentable, but is certainly understandable and, in certain situations, even quite reasonable. Up close, however, the rainbow fades. It becomes difficult to determine where the edges of halo are; where it exists and where it does not. Its central prerequisite, that changes in affect cause shifts in perception, becomes hopelessly confounded with the alternative explanation, that shifts in perception cause changes in affect. Various definitions of halo, while possessing a degree of face validity, do not allow an unequivocal measure of the strength of the bias due to halo or even its direction. In this study, the use of attributes which have clear physical counterparts, such as size or temperature, allows a relatively independent measure or bias due to halo. It then becomes possible to treat halo as a variable that can be explained by other variables. This begins to move halo effect from the status of a vague explanation to an operationally defined theoretical construct.

Halo has been defined in many ways. These are summarized in Table 1. Much of the work is by psychologists working with personnel evaluation forms. There it is found that a general attitude towards the person being rated, the ratee, results in a corresponding bias in the rating of more objective attributes, such as rated ability to perform a certain task. The measures of halo used have been of four kinds. The first is related to the correlation between attributes, the second to the variance of those attribute ratings for a given subject, the third as an additive interaction between the rater and the object, and the final measure is related to the degree to which preference predicts attribute judgments.

Defining halo as the excessive correlation between attributes goes back to the initial conceptualization of the concept (Wells, 1907; Thurstone, 1920). To the extent that preference causes attribute judgments, it will result in a spurious correlation between the attributes due to a preference factor. Unfortunately, as is pointed out by Bingham (1939), there is a degree of correlation expected even without the halo effect being operative. For example, it is reasonable to expect, for most subjects, that cornering ability and sportiness in automobiles correlate with preference regardless of halo. The problem of determining when the level of correlation becomes excessive has not, and probably cannot, be answered. Taylor and Hastman (1956) use average attribute correlation as a measure of halo to see if different preliminary instructions to raters reduce halo. While instructions as to the pitfalls of halo do appear to reduce the average correlation

TABLE 1
Summary of Operational Definitions of Halo

Definition	Defined On	References	Problems
1. Extra correlation between attribute ratings due to preference factor.	Rater	Wells (1907) Thurstone (1920) Bingham (1939) Koltov (1962) Taylor & Hastman (1956) Keaveny & McGann (1975)	Some correlation is expected. There is no way to determine how much indicates halo.
2. Lack of variance of attributes within object.	Rater and object	Brown (1968) Borman (1975)	Some variance is expected. No way to determine how little indicates halo. Primarily measures subject attention.
3. Additive rater x object interaction.	Rater and object	Guilford (1954)	Attributes must be scaled such that more is better. Contamination effect of factors other than preference.
4. Coefficient of preference predicting belief using average belief as a covariate.	Rater and attribute beliefs	Beckwith & Lehmann (1975)	Assumes average belief is halo free and represents the ultimately "true" belief for the individual.

between attributes, it is impossible to determine whether halo has been reduced or whether raters have merely learned to make artificially independent ratings.

Brown's (1968) definition of halo in terms of attribute variance is based on the notion that a given overall rating will produce correspondingly high or low ratings on the components. If the attributes are coded so that they all produce affective response in the same direction (e.g., more is better), then halo is hypothesized to reduce the variance of the attribute judgments about any given subject. Once again, however, there is no way to tell when the variance is sufficiently high to indicate lack of halo. By pointing out to raters the problem of low variance, Brown (1968) and Borman (1975) are able to significantly increase attribute variance after training, but it is unreasonable to conclude that the change represents less halo effect rather than an artificial scale effect by compliant subjects.

Working across raters, Guilford (1954) defines halo as a rater x object interaction term. By this definition, the preference of a rater for a particular object is

hypothesized to result in a constant added to or subtracted from each attribute for that object. Thus, if a particular rater likes an object, ratings on all dimensions should increase for that subject. There are two things to note about this formulation. First, as with the variance measure, all attributes must be coded in the same direction. Given halo bias, preference for an object will increase ratings on attributes where "more is better", decreasing those where "more is worse". The effect of two oppositely coded attributes tends to lessen or cancel Guilford's interaction term. The second problem is conceptual. We have here defined halo to be the effect of preference on attribute ratings. But preference need not be the only cause of a rater x object interaction. Attributes can themselves have a degree of intercorrelation so that a high rating on one attribute may lead to distorted ratings on other attributes. For example, in rating candidates, if one is considered to be intelligent, he is also likely to be rated higher on many other ratings, thus resulting in a rater x object interaction. To the extent that there are factors other than preference contaminating judgments, Guilford's measure will overestimate the halo effect.

There are, however, two advantages with this formulation of halo. First, even though significant rater x object interaction does not positively indicate the presence of halo, a lack of interaction argues strongly for the lack of halo effect. Second, the method provides a reasonable measure of halo defined in the larger sense of the biasing effect of all attributes, including preference, on each other. This, in contrast to either the correlation or the variance definitions, provides an independent measure of halo which can then be used as a basis to begin to understand the phenomenon.

Beckwith and Lehmann (1975) provide a definition of halo that goes even further towards making the concept operational. While it is not totally clear what their measure is, halo appears to be defined as the coefficient of preference in a regression predicting attribute ratings once the effect of average belief has been removed. This is:

$$B_{ij} = \delta_{j1} B_{ij}^* + \delta_{j2} A_i \quad (1)$$

Where:

B_{ij} = individual's attribute rating of stimulus i by attribute j .

B_{ij}^* = average value across all individuals in study of B_{ij} .

δ_{j1} = effect of average belief on individual beliefs.

δ_{j2} = halo effect of individual preference on beliefs.

A_i = attitude or preference toward object i .

Equation (1) is parameterized across objects rated by a given subject on an attribute. The idea here is that individual belief can be decomposed into two additive components, the first component representing the effect of average beliefs, which is assumed to be relatively halo free; and the second component representing the effect of preference which alters perception through halo. The larger δ_{j2} , the greater halo is presumed to be for that subject.

Beckwith and Lehmann use a system of simultaneous equations and two-stage least squares to estimate the parameters in their system, which, while commendable in principle, resulted in an indeterminate solution. In the case of the Beckwith and Lehmann paradigm, it appears that the coefficients of simple least squares estimation of Equation (1) result in very similar results as the more complex simultaneous system. Thus,

this definition of halo appears, in this instance, to be relatively independent of whether an individual or simultaneous system of equations is used.

Even taking the more simplified form of their model, there are two problems with measuring halo as in Equation (1). The first problem stems from the use of average belief as a surrogate to "true" belief. In doing so the model biases downward the effect of halo when there is considerable agreement across subjects with respect to preference and biases it upward when individual beliefs have a large, but legitimate, individual component.

For example, if different raters have similar tastes in objects, then average beliefs are likely to be contaminated by a kind of average halo effect. Thus, if all raters like a candidate, then it is reasonable that all attribute beliefs will be biased in a direction consistent with this average affect. In this case, the coefficient of A_i in Equation (1) will be attenuated since average attribute measures already contain a large component of individual preference. Thus, the effect of similarity of preference across subjects is to bias downward the measure of halo.

The opposite bias occurs when an attribute has a large, but legitimate, individual component. Consider the attribute of "fit" as it is applied to the preference for a jacket for an individual. To the extent that individual body sizes differ, the average measure of fit of any given jacket is likely to be relatively unrelated to individual preference. A large man will not have a strong preference toward a jacket that has the highest average fit. By contrast, individual fit is likely to be highly correlated with preference and thus lead one to conclude erroneously that there is a strong halo effect with respect to that attribute.

The final difficulty with the Beckwith and Lehmann formulation is making behavioral sense of Equation (1). That is, individual belief is formulated as a weighted combination of average belief and individual preference. But, following a perceptive comment by Johansson *et al.* (1976), it is hard to see how B_{ij}^* can be considered an input to an individual's belief when it is generally not known to the individual. Thus it is nonsensical to assert that Equation (1) represents what an individual could be doing, consciously or unconsciously. Such an average belief might make sense if the individual were asked what he felt the average belief was. Such a solution, however, runs into problems with respect to the direction of causality since a person's perception of average belief may be seriously tainted by their actual belief.

Thus, there are three problems with the Beckwith and Lehmann formulation of halo. The first concerns the identification problems arising from the use of a simultaneous system of equations, the second concerns the bias in halo estimates brought about by the likelihood that average belief is itself contaminated by halo, and the third concerns the behavioral meaningfulness of Equation (1) given that average belief is not typically known to the person making the rating.

A New Concept of Halo

In what follows, a concept of halo is developed that seeks to alleviate or avoid the above problems. Simultaneity is avoided by having all measures simply be correlation coefficients. The use of the average subjects' belief is avoided by using attributes that have clearly defined physical counterparts. This objective measure is then used in place of the average belief. Finally, the problem of using an unknown input in the

belief equation is avoided by defining halo as a function of the extent to which the objective value differs from the individual's belief. Thus the "true" value only enters to define perceptual error and as such is not expected to be known by the individual.

To illustrate the development, data from a study of 14 masters students with respect to their preferences toward 17 cities as sites for their first job is used. They were shown pairs of cities and asked which they would prefer to live in given that work related variables, such as opportunity for advancement, were kept constant. They were then asked how much more they would have to be paid per year to choose the less preferred city. From these dollar differences between pairs of cities, a one-dimensional preference scale was constructed for each subject using a procedure developed by Pessemier and Teach (1966) and described by Huber and James (1976). The process resulted in reliable individual preference scales of interval quality.

Three attributes were chosen for analysis, perceived size, average temperature, and opportunities for spectator sports in the city. These were measured on 10-point Likert Scales. The physical attributes are, respectively, population, average temperature, and the number of major professional sports teams in the city. Since the analysis is made on the basis of univariate relationships, no need was felt to include all determinant attributes. In this case, the attributes were chosen because it was relatively easy to find corresponding physical attributes that could serve as the basis for measuring perceptual error.

Thus the input to the measure of halo contains three elements for each individual: a preference or attitude score for each city, a set of beliefs for each city across three attributes, and finally a set of physical measures that correspond to the beliefs. The particular terms used and their operational definitions are provided in Table 2. Conceptually, halo is defined as the

TABLE 2
Summary of Concepts and Definitions

Concepts	Definitions
A_i Preference	Overall measure of affect towards stimulus i . Standardized within subject.
B_{ij} Psychological Attribute Rating	Perceived measure on stimulus i by attribute j . Standardized for each subject and attribute.
B_{ij}^t Physical Attribute Rating	Physical measure on stimulus i by attribute j . Standardized within attribute.
Psychophysical Transform	Best fitting linear transform predicting psychological attributes given their physical measures.
$B_{ij} = \alpha_j B_{ij} + u_{ij}$	
u_{ij} Perceptual Error	Deviation from perceived attribute of best fitting psychophysical transform.
Y_j Halo	Degree to which preference is related to perceptual error. Measured as the correlation between A_i and u_{ij} .

degree to which preference for an object biases beliefs about that object in such a way as to make them more consistent with the preference. Perceptual error is defined as the difference between the perceived rating of a city

and the best fitting linear approximation given its "true" value. If this residual is positive it means that relative to other cities the individual has judged this city to have a higher rating than is justified given its physical qualities. Halo is then simply the degree to which preference accounts for this perceptual error. This is operationalized as the correlation between the residual of the psychophysical transform and the preference score.

An example may make the logic of this operationalization more clear. Suppose that within the range of cities tested that a person likes warm climates--that being in a warm climate is important. Under halo it is hypothesized that if a person likes a city it will be misperceived to be warmer than it actually is. If correspondingly, less preferred cities are seen to be colder than is justified, then there will be a high positive correlation between perceptual error and preference.

Notice that the direction and the size of the halo coefficient will be a function of whether the attribute is liked and how much it is liked. In the example, if warmth is disliked then those cities that are disliked will be perceived as colder--producing a negative relationship. In a similar way, the importance of an attribute should be related to the actual size of the halo measure. Thus if warmth, or coldness, is very important to an individual then there is more pressure to bring beliefs in line with preferences. One measure of the degree of importance is the simple correlation between preference scores and beliefs. This measure reflects whether in general a high level of preference is associated with a high level of belief.

These two measures, the halo measure and the correlational importance measure are provided in Table 3. The

TABLE 3
Measures of Attribute Importance* and Halo for Fourteen Subjects

Attribute of City Subject	Size		Opportunities for Spectator Sports		Average Temperature	
	Imp.	Halo	Imp.	Halo	Imp.	Halo
1	-.15	-.24	.09	.13	-.66	-.37
2	.59	.52	.44	.35	.74	.28
3	-.29	-.17	.00	.30	-.11	-.17
4	.40	-.23	.50	.70	.33	.22
5	-.18	-.04	.05	.21	-.06	.03
6	.10	-.09	.09	-.04	.01	-.19
7	.78	.33	.85	.67	.16	.44
8	-.38	-.07	-.28	-.25	-.37	.10
9	.56	.44	.33	.30	.62	.29
10	.82	.82	.17	.42	.68	.40
11	.31	.08	.58	.61	.31	.41
12	-.04	-.05	-.21	.24	.41	.38
13	-.47	-.08	-.50	-.30	.39	-.04
14	-.13	-.53	.30	.33	-.82	-.44
Correlation Across Objects	.76 p < .001		.87 p < .001		.83 p < .001	

*Importance (Imp.) is measured as the simple correlation between attribute ratings and preference.

correlations across subjects are greater than .75 for each attribute. Moreover, it is visually clear that even within subjects the two measures covary strongly. Although this relationship is in part due to a computational artifact stemming from the use of preference in both measures, the strength of the relationship makes

behavioral sense and lends credence to this measure of halo.

The relative importance of preference in determining attributes can be looked at from another perspective. For the three attributes, size, spectator sports and warmth, the physical values of these attributes accounted for an average of 61%, 60% and 78% of their variance, respectively. From values of halo in Table 3, it is evident that preference accounts for about 10% of the remaining variance. Thus, halo accounts for a relatively small proportion of the error in attribute judgments. This result, for attributes that have sharply defined physical counterparts, differs from the result of Beckwith and Lehmann study, using more subjective attributes. They found that preference accounts for about 30% of the variance in attribute judgments. This makes sense, the more subjective the attribute, the easier it is for preference to legitimately distort judged attributes. By contrast, the objectivity of a variable like size, or average warmth, makes it more difficult for subjects to distort its value.

Thus, this measure of halo enables one to estimate halo in a way that makes conceptual sense and is simple computationally; for reasons of parsimony alone it is to be preferred to the Beckwith and Lehmann version.

It has two disadvantages. First as operationalized it can only be applied to attributes that have clearly defined physical counterparts. While the method could be used by simply substituting the average subjective belief for the objective measure, such a solution would lead to the problem of average belief contaminated by halo referred to earlier. A partial solution is to develop elaborate psycho-physical transforms that predict the subjective attribute (say, good dating opportunities for a city) as a function of several objective attributes (say, the percent of population that is 18-25 years old and the number of nightclubs per capita).

The second problem is more serious but cuts across all measures of halo presented thus far. This is the problem of the direction of causality. There is an alternative explanation for the results we have presented which we shall call the "random experiences explanation." It will be shown that this accounts for the same results as the halo explanation.

Consider once again the person who likes warm cities. Suppose he erroneously considers a city, say St. Louis, to be warmer than it really is. Given that there is a strong positive affective value attached to warmth, St. Louis can be expected to have greater preference due to the misperceived warmth. Furthermore, the degree of preference bias will be approximately proportional to the degree of positive affective value attached to coldness. Thus, the random experiences and halo explanations account for the same experimental results. In the data provided so far, they are empirically equivalent. The difference between the two is in the direction of causality between beliefs and preference. The correlational methods used cannot determine causal direction. Moreover, it is likely that neither explanation is true to the exclusion of the other, but rather that causality runs in both directions to a greater or lesser extent depending on the subject and the attribute. The next section provides a way to begin to identify the predominant direction of causal flow for different attributes and subjects.

Estimating Causal Direction

If the link between preference and perceptual error is primarily due to the translation of legitimate misperception of attributes into preference ratings, then one would expect this link to diminish as one becomes

more familiar with the objects being judged. On the other hand, to the extent that the halo explanation is correct, selective perception should operate to keep the relationship between preference and perceptual error high even as one becomes more familiar with the stimuli.

In the present study, subjects were asked to specify level of familiarity with each city on a ten-point scale. This can be used to test the halo versus the random experiences explanations by estimating the following equation.

$$u_{ij} = \gamma_j A_i$$

$$u_{ij} = (\gamma_{j1} + \gamma_{j2} \cdot \text{FAM}_i) A_i \quad (2)$$

$$u_{ij} = \gamma_{j1} \cdot A_i + \gamma_{j2} (\text{FAM} \cdot A_i) \quad (3)$$

Where:

u_{ij} = Perceptual error on city i by attribute j .

A_i = Preference for city i .

FAM_i = Familiarity with city i on a ten-point scale.

γ_j = Average effect of preference on error for attribute j , halo effect.

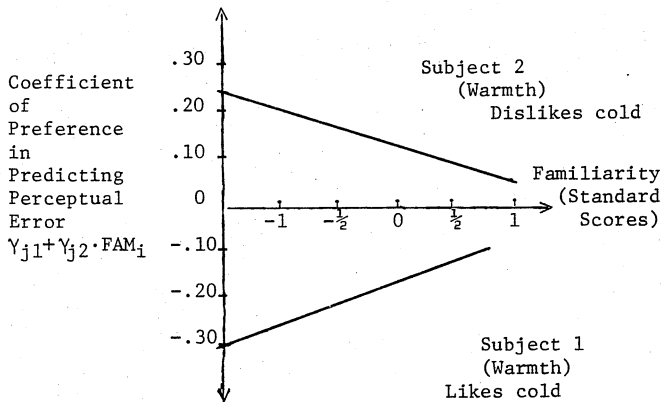
$\gamma_{j1} + \gamma_{j2} \cdot \text{FAM}_i$ = Halo effect taking account of the effect of familiarity.

What is expected under the random experiences explanation is that familiarity will reduce the correlation between preference and error. Thus γ_{j1} and γ_{j2} should have opposite signs. If halo is operant, familiarity should have no effect and the coefficient of γ_{j2} should be zero. Equation (2) can be estimated by changing its form to Equation (3) and simply using regression. This was done for each of the three attributes for the fourteen subjects in the study. Two-thirds of the cases tested did not have significant γ_{j2} terms at a 0.20 level of significance, thus generally supporting the halo hypothesis. Of course, there were only 17 cities in the regression so the power of the test was not very strong. In addition, the correlation between A_i and $\text{FAM}_i \cdot A_i$ results in multi-collinearity which further weakens the power of the test. Since, however, neither hypothesis is expected to be true to the exclusion of the other, the magnitude of the γ_{j2} term can be used to segment subjects on the basis of relative magnitude of the halo effect.

Figure 1 provides the familiarity-adjusted halo level for subjects displaying low halo effects. Notice that for subject 1, preference is negatively correlated with misperceptions of average temperature, but that the degree of this negative correlation diminishes on those cities about which the subject is more familiar. This implies that as cities become more familiar, the perceptual bias due to preference will diminish. Had halo effect been the explanation for the relationship between preference and error, then the slope of the line would have been more nearly horizontal.

Subject number 2 shows a high positive correlation between preference and affective error. This results from a general preference for warm climates so that preferred cities are misperceived in the direction of being warmer. Once again, however, note that the strength of this bias decreases with increased familiarity. In both of these cases, one would hypothesize that these subjects would be receptive to new information about the warmth of the

FIGURE 1
Effect of Familiarity on the Degree to Which Preference Correlates with Perceptual Error



city and it would be reflected in their preference for cities.

By contrast, where the slope is not significant, it can be inferred that preferences are the primary cause of beliefs and that new information is not likely to reduce the bias due to preference. This can be inferred from the fact that for those attributes where γ_{j2} is not significant, increasing familiarity does not reduce error.

In this study, approximately one-third of the attributes resulted in equations which contradicted the halo explanation at a .20 significance level. This implies that while halo is dominant in the majority of cases, there is a significant minority for which random experiences provides a better explanation. Therefore, it is important to make assessments of the relative importance of the two explanations on a case-by-case basis rather than attempting generalizations across all subjects or attributes. Seen in this way, the significance of γ_{j2} represents a descriptive statistic that enables one to group subjects in terms of the degree of appropriateness of the halo explanation. The ability to measure the relative importance of halo has promise as a marketing tool. To the extent that a subject's perceptions are contaminated with halo, it is unlikely that further communication, contrary to this preference, is going to be accepted in an unmodified form. On the other hand, to the extent that marketers can identify those whose perceptual bias is reduced with greater familiarity, it is likely that those individuals will be good targets for marketing communications.

Summary and Implications

In contrast to other definitions of halo, the operationalization presented here is simple and logically follows from the idea of halo as perceptual bias due to preference. While the issue of the direction of causality between preference and beliefs is not completely resolved, the use of familiarity ratings provides a promising technique to identify those cases where the causal flow is predominantly in one direction or the other.

The study suggests several paths for future research. While in this model the physical value is defined as the "true" value of an attribute, other measures are possible. Conceptually, what is needed is that the "true" value approximate the halo free rating the individual would give provided enough time and information. For those variables that have clearly defined physical counterparts, the physical values serve this function well. As mentioned earlier, the average rating across subjects is likely to fail on both criteria; it will

neither be halo free nor approximate the asymptotically true value for most individuals. To generalize the model, what could be done is to expand the psychophysical transform to include several physical components that make up higher-order psychological ratings. For example, multiple regression could be used to predict ratings on sportiness in automobiles as a function of such measurable quantities as width-to-height ratio, cornering ability, as well as dummy variables such as rack-and-pinion steering and disk brakes. Such extended psychophysical transforms would be useful not only in measuring halo, but also in providing guidance to designers attempting to create a sporty car.

Familiarity proved to be a promising variable to discriminate between the halo and the random experiences explanations. In this case, a global familiarity rating on each city was used rather than separate familiarity ratings on each attribute. This oversimplification could result in problems to the extent that subjects are more familiar with one aspect of a city than another. Thus, the model could be improved by including a familiarity or degree of certainty estimate for each attribute.

Finally, experimentation is needed to determine if in fact those who appear to have less halo in their judgments are more susceptible to persuasive communications. If this turns out to be true, then the methodology provided here could become a valuable basis for segmenting subjects in terms of sensitivity to advertising or new information.

References

- Neil Beckwith and Donald Lehmann, "The Importance of Halo Effects in Multi-Attribute Attitude Models," *Journal of Marketing Research*, 12(August, 1975), 265-275.
- W. V. Bingham, "Halo, Invalid and Valid," *Journal of Applied Psychology*, 23(1939), 221-228.
- Walter C. Borman, "Effects of Instructions to Avoid Halo Effect on Reliability and Validity of Performance Evaluation Ratings," *Journal of Applied Psychology*, 60 (October, 1975), 556-560.
- Eva Metzger Brown, "Influence of Training, Method, and Relationship on the Halo Effect," *Journal of Applied Psychology*, 52(June, 1968), 195-199.
- J. P. Guilford, *Psychometric Methods* (New York: McGraw Hill, 1954).
- Joel Huber and William James, "The Marginal Value of Physical Attributes: A Dollarmetric Approach," presented at the 7th American Marketing Association Attitude Conference, Hilton Head, S.C., February 1976.
- Johny K. Johansson, D. J. MacLachlan and R. F. Yalch, "Halo Effects in Multi-Attribute Models: Some Unresolved Issues," *Journal of Marketing Research*, 8(February, 1976), 414-417.
- T. J. Keaveny and A. F. McGann, "A Comparison of Behavioral Expectation Scales and Graphic Rating Scales," *Journal of Applied Psychology*, 60(December, 1975), 695-703.
- Barbara Koltov, "Some Characteristics of Intrajudge Trait Intercorrelations," *Psychological Monographs*, No. 552 (1962).
- E. A. Pessemier and R. D. Teach, "A Single Subject Scaling Model Using Judged Distances Between Pairs of Stimuli," Institute Paper No. 143, Krannert Graduate School of Industrial Admin., Purdue University (1966).

E. K. Taylor and R. Hastman, "Relation of Format and Administration to the Characteristics of Graphic Rating Scales," Personnel Psychology, 9(1956), 181-206.

HALO EFFECTS AND LOCATION PREFERENCES

William L. James, University of Alabama
Forrest S. Carter, Georgia Institute of Technology

Abstract

The two methods of measuring halo that have been used in marketing were compared and were found to be highly correlated. Some evidence was found to support the hypothesis that halo decreases as preference increases. Halo was not found to increase as familiarity decreases. Halo was found to be a more important source of error than familiarity. Attributes with less clearly defined physical correlates were found to have more halo than those attributes with clear physical correlates.

Introduction

The focus given to attitude research in marketing and the behavioral sciences has led to a renewed interest in halo effects. The halo effect was noticed as early as 1907, but was named by Thorndike (1920). There have been two slightly different ways in which halo has been conceptualized. Tiffin and McCormick (1965) in their classic textbook view halo as the domination of all traits by a particular trait. A more common view is that halo occurs when an overall impression, such as preference, dominates the traits being rated. This latter view appears to dominate in the marketing literature. The concept of halo is important to marketing researchers due to its effect on the prediction of preference, the development of product spaces, and the identification of marketing opportunities. Any time that attribute ratings are collected the researcher must consider the potential impact of halo on the relationships being investigated. Halo might be beneficial if the objective is to predict preference from attribute ratings since it will increase the correlation between the attributes and preference. Halo is, however, detrimental when attribute scores are used to develop product spaces or identify marketing opportunities. It would be a mistake to develop or reposition a product to fit some set of consumer ratings that are heavily influenced by halo. Thus, it is necessary to measure the degree to which attribute ratings are influenced by halo in a given situation. If marketers could identify those consumers who exhibit high levels of halo in attribute judgements it might be possible to determine their susceptibility to advertising and other forms of marketing activity.

Background and Perspective

The first empirical research on measuring halo appears to have been done by Symonds (1925). He used partial correlations between traits to demonstrate that halo raised the correlations of trait ratings by .245 on the average. He speculated that large halo in a trait occurs if a trait is not easily observed or if it is not clearly defined. This line of research led to the view of halo as the excessive correlation between attributes. Unfortunately, as Bingham (1939) points out, some correlation between attributes is to be expected, and there seems to be no clear cut method to determine when the correlation is excessive. This approach usually led to factor analytical solutions where the first factor extracted was assumed to be a measure of overall attitude. The trait intercorrelations were then studied after this factor had been removed. Later research using this definition, such as

Keaveny and McGann (1975), focused on average correlations between attributes and attempted to lower these intercorrelations through instructions and training of raters.

A second approach to the measurement of halo is due to Guilford (1954) who viewed halo as an interaction between rater and object. Guilford (1954) defined several different types of errors which can occur when human raters are used. These errors are halo, leniency error, logical error and rater-trait error. Leniency error is defined as the tendency of a subject to overvalue or undervalue objects in general. Logical error occurs when subjects give similar ratings to objects on traits that appear to the subject to be closely related. Rater-trait error is defined as the tendency of a subject to overvalue or undervalue an object on a particular trait due to a favorable or unfavorable attitude towards the object. An analysis of variance design is then used to test for the rater by object interaction. This definition will overestimate halo to the extent that there are factors other than preference contaminating attribute judgements. An advantage of this approach is that a lack of a significant interaction term indicates the absence of halo.

A third approach to the measurement of halo, used by Brown (1968), focuses on the variance of ratings of an object across attributes. The higher the variance in ratings for an object, the less halo possessed by the object. It is important to note that this approach assumes that all attributes for a given subject are scored so that the subject views more as better. Thus, this approach requires a judicious choice of attributes or the measurement of preference for each subject so that it can be determined whether any recoding is in order for the halo measure. Unfortunately this measure does not indicate how little variance is needed for halo to be present. Wilkie and McCann (1972), and Wilkie, McCann, and Riebstein (1973) were the first to apply this approach in a marketing context. They found, among other things, that instructions could increase variance, that higher variance was found for the most preferred brand than for the other brands, and suggested that low brand familiarity contributes to low variance in ratings. Koltuv (1962) found that as familiarity decreases, halo increases, and that as halo increases the relevance of the trait increases. This is the only approach which has been used by both psychologists and by marketers.

A fourth approach to the study of halo is a regression model by Beckwith and Lehmann (1975) which has only appeared in the marketing literature. Their general findings agree with those from the industrial psychology literature. The most important contribution of this model is that it allows the assessment of the degree of halo exhibited by individual attributes for each subject. This approach views beliefs as being composed of a true belief rating, estimated from average belief scores, and halo error, estimated using preference. The model for an individual subject is given below:

$$B_{ij} = b_1 \bar{B}_{ij} + b_2 P_i + u_{ij} \quad (1)$$

where: B_{ij} is the belief rating for object i on attribute j

\bar{B}_{ij} is the average belief rating for object i on attribute j
 P_i is the preference for object i
 b_{1j} is the importance of the true value of the object on attribute j
 b_2 is the measure of degree of halo for the object on attribute j
 u_{ij} is random error
 All variables are standardized.

Purpose of the Study

The present study follows up on the Beckwith and Lehmann (1975) and the variance approaches. Specifically this paper investigates four areas:

- (i) The relation of the variance measure to both familiarity and preference within a subject.
- (ii) The convergence of the variance approach and the Beckwith and Lehmann (1975) approach.
- (iii) Whether halo is greater for attributes with less well defined physical correlates.
- (iv) Whether halo effects can be separated from misperceptions due to a lack of familiarity.

Data Base

Fourteen masters students rated seventeen cities on nine attributes plus familiarity. Then preference for the cities as places to work was measured using a technique described by Huber and James (1976), resulting in dollarmetric scale values of preference. The attributes were measured on 10 point Likert scales. The cities were selected from a list of 40 used in a pretest. The attributes and the dollarmetric scale were also pretested. In the variance analyses all attribute ratings were rescaled whenever necessary so that for each subject all attributes had positive correlations with preference. The nine attributes rated were: population; spectator sports; opportunities for indoor recreation; opportunities for outdoor recreation; cultural activities; warmth of the climate; pleasantness of the summer; pleasantness of the winter; and snowfall.

Results

- (i) Within subject analyses using the variance measure are presented in Table 1:

TABLE 1

Subject	Preference vs. Familiarity	Preference vs. Variance	Familiarity vs. Variance
1	.46*	.17	.26
2	.28	-.45*	-.18
3	-.22	-.10	.15
4	.42*	.59**	.43*
5	.27	.35	.20
6	-.33	.18	-.31
7	.74***	-.07	-.22
8	-.21	-.09	-.13
9	.33	.57**	.72***
10	.10	-.19	-.39
11	-.07	.10	-.10
12	.41	.42*	.34
13	-.13	-.08	.13
14	.44*	.19	.33

*significant at .10 level (2 tailed test)
 **significant at .05 level (2 tailed test)
 ***significant at .001 level (2 tailed test)

Since there are 14 subjects and only a few have significant correlations a reasonable way to test whether relationships exist is to perform a binomial test. Wilkie, McCann, and Reibstein (1973) found that the lowest preference had the lowest variance, therefore there are 3 correlations out of 14 that are significant at the .05 level for a one tailed test. The probability of 3 or more successes out of 14, given the probability of an individual success is .05, is .0300. Thus, the results are highly unlikely to be random perturbations. Rather, it is reasonable to conclude that a positive relationship exists between preference and variance. To test whether preference and familiarity are positively correlated we find that there are 4 significant positive correlations at the .05 level for a one tailed test. The probability of 4 successes out of 14 is only .0041. Thus it is unlikely that these came about by chance, and it can be concluded that preference and familiarity are positively correlated. To test whether halo is due to misperceptions of the product due to a lack of familiarity we note that there are 2 significant positive correlations at the .05 level for a one tailed test. The probability of 2 successes out of 14 is .1529, and is thus, likely to have occurred on a chance basis alone. Consequently the table does not offer any evidence that for this product there is a relationship between halo and familiarity.

- (ii) The degree of convergence between the variance approach and the Beckwith and Lehmann (1975) approach was assessed by converting the halo measures, b_2 in equation (1), into an overall measure for each subject. The overall halo measure was simply the sum of the absolute values of the individual b_2 . This measure of overall subject halo was then correlated with both the mean variance of a subject's ratings, and with the variance of the variance measure. The correlation of halo with the mean variance was $-.46$ ($p = .095$) and is in the expected direction. The variance of the variance measure was expected to be a better measure since it measured the dispersion rather than the absolute level of variance in product ratings. This variance of the variance measure was correlated $-.65$ ($p = .011$) with halo, and is in the expected direction. Thus, the two different halo measures agree and provide convergent measures of halo.

- (iii) In order to test whether those variables with less clearly defined physical correlates exhibited greater halo, the Beckwith and Lehmann (1975) measure was used. Table 2 shows the mean and variance for halo scores by attribute and by grouping:

TABLE 2

Attribute	Mean Halo	Variance of Halo
Population	.16	.04
Spectator Sports	.16	.01
Warmth of Climate	.14	.01
Snowfall	.15	.02
Indoor Recreation	.35	.06
Outdoor Recreation	.37	.06
Cultural Activities	.26	.04
Pleasantness of Summer	.39	.05
Pleasantness of Winter	.16	.01
4 Physical Attributes	.15	.02
4 Less Physical Attributes	.34	.06
5 Less Physical Attributes	.30	.05

Population, spectator sports, warmth of climate, and the amount of snowfall have clear physical correlates and were pooled for this analysis. Indoor recreation, outdoor recreation, cultural activities, and pleasantness of the summer were pooled as attributes with less

clearly defined physical correlates. When these two groups are compared, the less physical attributes had significantly higher halo scores ($t = 5.03$) at the .005 level. If pleasantness of the winter is pooled with the less physical attributes, the difference ($t = 4.26$) is still significant at the .005 level.

(iv) In an attempt to separate halo from perceptual error, familiarity was added to equation (1). It is assumed that familiarity reflects possible misperceptions due to random nonrepresentative experiences with the object. Table 3 shows the beta weights for a typical subject.

TABLE 3

Halo Versus Random Experiences Explanation

Attribute	Average Score		Preference		Familiarity	
	b	R	b	R	b	R
Population	.83	.96	-.29	-.14	.23	.45
Spectator Sports	.61	.81	---	.35	.42	.77
Indoor Recreation	.44	.22	.55	.88	---	.39
Outdoor Recreation	.29	-.23	.70	.44	---	.39
Cultural Activities	.54	.50	.54	.81	---	.40
Warmth of Climate	-.74	.22	-.32	-.84	---	-.31
Pleasantness of						
Summer	---	-.32	.75	.94	---	.17
Pleasantness of						
Winter	.96	.31	---	-.62	.14	-.08
Snowfall	.62	-.17	.45	.84	---	.27

b = beta weight, listed only if significant at .10
R = structural coefficient

The analysis indicates that familiarity is usually less important than halo as a contributor of error, and that for this product it is not tenable to attribute most of the halo to a lack of familiarity with the cities in the questionnaire.

Conclusion

The most important conclusion is the finding that both the variance measure, used fairly extensively in psychology, and the Beckwith and Lehmann (1975) measure correlate highly. The finding that halo was higher for variables with less clearly defined physical correlates implies that in order to change attitudes prior to trial of a product it may be necessary to focus on those attributes with clearly physical correlates. The low amount of halo found in this study indicates that there is wide variation in the amount of halo in different types of products, and that until more is known about which types of products are heavily influenced by halo each product will have to be investigated on its own.

References

Neil Beckwith and Donald Lehmann, "The Importance of Halo Effects in Multi-Attribute Attitude Models," Journal of Marketing Research, 12 (August, 1975), 265-275.

W. V. Bingham, "Halo, Invalid and Valid," Journal of Applied Psychology, 23, (1939), 221-228.

Eva Brown, "Influence of Training, Method, and Relationship on the Halo Effect," Journal of Applied Psychology, 52, (June, 1968), 195-199.

J. P. Guilford, Psychometric Methods. New York, McGraw-Hill, 2nd edition, 1954.

Joel Huber and Bill James, "The Monetary Worth of Physical Attributes: A Dollarmetric Approach," paper presented at the Seventh Annual Attitude Research Conference, Hilton Head, South Carolina, on February 12th, 1976.

Timothy Keaveny and Anthony McGann, "A Comparison of Behavioral Expectation Scales and Graphic Ratings Scales," Journal of Applied Psychology, 60, (December, 1975), 695-703.

Barbara Koltuv, "Some Characteristics of Intrajudge Trait Intercorrelations," Psychological Monographs, 76, Whole No. 552, (no. 33, 1962).

Percival Symonds, "Notes on Ratings," Journal of Applied Psychology, 9, (1925), 188-195.

E. L. Thorndike, "A Consistent Error in Psychological Ratings," Journal of Applied Psychology, 4, (January, 1920), 25-29.

Joseph Tiffin and Ernest J. McCormick, Industrial Psychology, Prentice-Hall, Englewood Cliffs, New Jersey, 5th edition, 1965.

William Wilkie and John McCann, "The Halo Effect and Related Issues in Multi-Attribute Attitude Models - An Experiment," Paper No. 377, Institute for Research in the Behavioral, Economic and Management Sciences, Krannert Graduate School of Industrial Administration, Purdue University, October, 1972.

William Wilkie, John McCann and David Reibstein, "Halo Effects in Brand Belief Measurement: Implications for Attitude Model Development," Proceedings, Fourth Annual Conference, Association for Consumer Research, November, 1973.

ECONOMETRIC ESTIMATION OF HALO EFFECT:
SINGLE VS SIMULTANEOUS EQUATION MODELS

Albert C. Bemmaor (doctoral candidate), Purdue University
Joel C. Huber, Columbia University

Abstract

This paper analyzes a single equation model of belief as a function of preference and compares it to a simultaneous equation model. Using statistical tests, this study shows that the single equation model is well specified whereas the simultaneous equation model leads to statistically significant differences in the estimates.

The problem of measurement of halo effect has been extensively dealt with in the marketing literature. In a recent paper, Huber and James present a review of the different operational definitions of halo (Huber and James, 1976). Beckwith and Lehmann use econometric methodology to assess the halo effect (Beckwith and Lehmann, 1975). They analyze individual data on attitudes and beliefs towards television shows. After estimating a single equation model of the effect of beliefs on attitude, they use a simultaneous equation model to eliminate the "intertangled effects" of attitude and beliefs. Their study has given rise to criticism and further research in this area (Wittink and Ness, 1976; Johansson, MacLachlan and Yalch, 1976). The purpose of this paper is to test the Beckwith and Lehmann single equation model of the effect of beliefs on attitude for specification errors, and evaluate the implications of a misspecified model. In the first section, the Beckwith and Lehmann model is described. The second section briefly presents the specification error tests. The third section summarizes the results of these tests. In the last section, we assess the effect of a misspecified model on the parameter estimates.

Beckwith and Lehmann Simultaneous Equation
Model between Attitude and Beliefs

The Beckwith and Lehmann simultaneous attitude-belief model can be represented as follows:

$$A_i = \sum_{j=1}^n w_j B_{ij} + \gamma A_i^* + \mu_0 \quad (1)$$

where

- A_i = individual's attitude towards stimulus i
- B_{ij} = individual's belief towards stimulus i on attribute j
- A_i^* = average attitude of all persons towards the stimulus i
- n = number of attributes

and where the unknown parameters to be estimated are:

- w_j = weight of attribute j
- γ = weight of average attitude A_i^*
- μ_0 being a random disturbance.

The belief equations can be represented as follows:

$$B_{ij} = \beta_j A_i + \gamma_j B_{ij}^* + \mu_j \quad \text{for } j=1, \dots, n \quad (2)$$

where

B_{ij}^* = average belief about stimulus i on attribute j
the other variables being previously defined, and where the unknown parameters to be estimated are:

- β_j = importance of attitude in predicting belief B_{ij} or halo effect
- γ_j = importance of average belief B_{ij}^* to belief B_{ij}

In this paper, we focus our attention on equation (2), the question being: Is (2) well specified? Is it sufficient to estimate halo, or do we need to add (1)? To some extent, we are asking the problem in a direction opposite from that as posed by Beckwith and Lehmann since they consider whether equation (2) should be added to equation (1) in order to estimate the importance weights w_j . However, here, we are not concerned with the estimation of importance weights, but rather concerned with the estimation of halo. We will attempt to show that equation (2) is well specified and therefore is adequate to estimate halo without (1).

Specification Error Tests

The tests which are used have recently been developed by Ramsey (Ramsey, 1969; Ramsey, 1974). These tests detect the following kinds of specification errors: omitted variable, incorrect functional form, simultaneous equation problem, heteroskedasticity and normality of the disturbances. Consider the model:

$$y = X\beta + \mu \quad (3)$$

where y is an $N \times 1$ vector, X is an $N \times K$ matrix of rank K , β is a $K \times 1$ vector of coefficients and μ is an $N \times 1$ vector of disturbances. Ramsey denotes the Full Ideal Conditions assumed when using ordinary least squares estimation method on the model above as: i) the disturbance terms are normally distributed with mean zero and constant variance, and are independent, ii) they are independent of the explanatory variables, iii) the matrix $X'X$ is nonsingular.

In the case of misspecification of (3), these Full Ideal Conditions are violated. The tests consist in analyzing the residuals of the model in order to detect such errors. The five errors considered are broken down into three groups. Group A errors consist of omitted variable, incorrect functional form and the simultaneous equation problem. These errors lead to failure to fulfill the assumption that the mean of the disturbances is zero. Group B errors are those which lead to the violation of the assumption of constant variance of the disturbances, an example of which is heteroskedasticity. Group C errors are those in which the disturbances are not normally distributed so that the distribution of the regression coefficients is affected.

Three tests are used for the analysis: RESET (having a central F distribution), BAMSET (based on a central chi-square distribution) and WSET (whose critical bounds have been defined by Shapiro and Wilk (Shapiro and Wilk, 1965)). RESET is designed to detect Group A errors, BAMSET is used for the detection of Group B errors and WSET for the detection of Group C errors. Further details on these tests are provided by Ramsey (Ramsey,

1974). Ramsey and Gilbert suggest that RESET and BAMSET seem to be independent under BLUS residuals. On the other hand, BAMSET and WSET are not independently distributed under BLUS residuals but do appear so under OLS residuals (Ramsey and Gilbert, 1969).

Under the Full Ideal Conditions, the ordinary least squares residuals are distributed as $N(0, \sigma^2 M)$, i.e., they do not have a scalar covariance matrix. Therefore, an alternative estimator of the vector of disturbances μ can be used to detect specification errors via the analysis of residuals. The BLUS residuals $\tilde{\mu}$ (Best Linear Unbiased Scalar covariance matrix residuals) developed by Theil are such that $\tilde{\mu} \sim N(0, \sigma^2 I_{N-K})$ under the Full Ideal Conditions (Theil, 1965). Ramsey and Gilbert showed that for the case of omitted variable, incorrect functional form and simultaneous equation problem (Group A errors), the use of $y = X\beta + \mu$, $\mu \sim N(0, \sigma^2 I)$ as the true regression model leads to a BLUS residual vector distributed as $N(A'\tilde{\epsilon}, \sigma^2 I_{N-K})$ where $\tilde{\epsilon}$ is a nonstochastic vector the definition of which depends upon the particular misspecification, and A is an $(N-K) \times K$ matrix such that $A'X=0$, $A'A=I_{N-K}$ and $AA'=I-X(X'X)^{-1}X'=M$ (Ramsey and Gilbert, 1969). In the case of heteroskedasticity (Group B errors), i.e. $\mu \sim N(0, \sigma^2 \Omega)$, the distribution of $\tilde{\mu}$ derived from the use of $y = X\beta + \mu$, $\mu \sim N(0, \sigma^2 \Omega)$ as the true model, is $N(0, \sigma^2 A'\Omega A)$. Finally, in the case of nonnormality (Group C errors), the BLUS residuals are such that $\tilde{\mu} \sim D(0, \sigma^2 I_{N-K})$, D being any distribution other than the normal distribution. Comparing the performances of the test statistics using BLUS and OLS residuals, Ramsey and Gilbert recommend the use of BLUS residuals when the primary concern is about Group A errors (Ramsey and Gilbert, 1972). As we are mostly concerned about the problem of simultaneity between belief and attitude, BLUS residuals are used in this study.

Test Results

The data used in this study were collected by Huber and James (Huber and James, 1976). Individual analyses were made on the preferences of fourteen master students towards seventeen U. S. cities as sites for their first jobs. They were shown pairs of cities and asked in which they would prefer to live, given that work related variables such as opportunity for advancement were kept constant. A dollarmetric scale was developed for each individual and used as a measure of preference (Pessemier and Teach, 1966; Huber and James, 1976). Three attributes were chosen for analysis: size of the city, warmth of the climate and opportunities for spectator sports. These attributes were measured on ten-point Likert scales. The results of the tests for the first attribute "size of the city" are presented in Table 1. This table indicates the

TABLE 1
Ramsey test results for the Beckwith and Lehmann model* (size of the city)

	RESET	BAMSET	WSET	DECISION
Individual 1	.4246	.8642	.9424	
Individual 2	.6160	1.9717	.9777	
Individual 3	.7040	.3371	.9322	
Individual 4	.3365	1.2750	.8536 (r)	R
Individual 5	1.9265	1.9725	.9814	
Individual 6	1.3585	.1867	.9304	
Individual 7	1.6341	.8947	.9501	
Individual 8	.1594	.4891	.9360	
Individual 9	2.2963	1.5501	.9275	
Individual 10	.5609	3.0402	.9755	
Individual 11	.6131	3.2421	.9674	
Individual 12	.8818	.2661	.9597	
Individual 13	1.5896	1.9056	.9295	
Individual 14	.0266	.2864	.9572	

Number of rejections (by column) 0 0 1 1

* An r indicates rejection of the null hypothesis of no specification error at the .1 level for a given test. An R in the column labeled DECISION indicates rejection of the hypothesis of no specification error for a given individual using model I because one or more of the test statistics rejected the null hypothesis.

number of times the null hypothesis of no specification error was rejected at an $\alpha = .10$. For each individual, the model is rejected when the null hypothesis of no specification error is rejected by one or more of the tests. These results are indicated in the column DECISION. For the first attribute, the null hypothesis is rejected only once (in favor of the alternative hypothesis of Group C errors). It should be noted that the null hypothesis is never rejected in favor of the hypothesis of Group A errors. Table 2 presents the test

TABLE 2
Overall performance of the Beckwith and Lehmann model for all three attributes

	Number of Rejections of H_0		
	in favor of Group A errors (RESET)	in favor of Group B errors (BAMSET)	in favor of Group C errors (WSET)
Size of the city	0	0	1
Warmth of the climate	2	1	2
Sports Opportunities	0	1	1
Proportion of rejections	2/42	2/42 2/42	4/42

results for all three attributes. In total, the null hypothesis of no specification error is rejected only twice against the alternative hypothesis of omitted variable, incorrect functional form and simultaneous equation problem (Group A errors). Also, for Group B errors, the Beckwith and Lehmann model is misspecified in two cases only. The assumption of normality of the residuals, tested by WSET, seems to be the most critical one, as evidenced by the number of rejections of the null hypothesis. Table 3 indicates the number of rejections

TABLE 3
Individual and model compatibility for all three attributes*

	Size of the city	Warmth of the climate	Sports Opportunities
Individual 1			
Individual 2			
Individual 3		R	
Individual 4	R	R	
Individual 5		R	
Individual 6			
Individual 7			
Individual 8			
Individual 9			R
Individual 10			
Individual 11			
Individual 12			R
Individual 13			
Individual 14		R	

* An R in any given column indicates that the specified model was rejected when applied to the indicated individual.

for each individual. Except for Individual 4, the Beckwith and Lehmann model seems to be appropriate for at least two attributes out of three. The major result of this analysis is that equation (2) appears to be well specified for most individuals. Despite the fact that

the simultaneous equation model might make theoretical sense, it is contradicted by the empirical evidence with this set of data.

A Monte Carlo analysis of the small sample properties of the tests discussed here is reported in Ramsey and Gilbert (Ramsey and Gilbert, 1972). The basic conclusion is that RESET and BAMSET are reasonably powerful tests against their respective alternatives. The small sample properties of WSET have also been examined by Monte Carlo analysis (Ramsey and Gilbert, 1969). The tests are capable of differentiating between Group A, Group B and Group C errors. However, it is not possible to discriminate between errors within a given group. Hence, the rejection of the model by RESET for two individuals (see Table 2) may not necessarily be due to the simultaneous equation problem but implies the presence of Group A errors in the model. On the other hand, acceptance of the null hypothesis against the hypothesis of a given group of errors implies that all the types of errors comprised in that group are excluded from the model.

Effect of a Misspecified Simultaneous Equation Model on the Parameter Estimates

Summarizing the findings of their study, Beckwith and Lehmann state: "The TSLS regression results indicated that ... for the six belief equations ... the regression coefficients (remained) relatively unchanged from OLS, on average" (Beckwith and Lehmann, 1976). Hence, their model specification does not seem to have much affected the estimates of the halo effect β_j and of γ_j . The TSLS estimates of β_j and γ_j were calculated and compared to the OLS estimates.¹ Table 4 presents these estimates

TABLE 4
Comparison of OLS and TSLS estimates of β and γ for "size of the city" (standard errors in parentheses)

	OLS β	TSLS β	OLS γ	TSLS γ
Individual 1	-.3107 ^a (.348)	.2199 (.489)	1.0627 (.254)	1.0234 (.276)
Individual 2	.1111 (.063)	.1383 (.078)	.5942 (.123)	.5702 (.130)
Individual 3	-.0600 ^a (.168)	.2311 (.302)	.8799 ^d (.209)	.9895 (.248)
Individual 4	-.0473 ^a (.351)	-2.1086 (2.486)	1.1612 ^a (.167)	1.6134 (.611)
Individual 5	-.0107 ^a (.083)	.1201 (.129)	1.0356 (.151)	1.0804 (.167)
Individual 6	-.0624 ^a (.073)	-.1695 (.100)	1.2722 (.189)	1.3369 (.206)
Individual 7	.0938 (.080)	.1064 (.139)	1.0572 (.177)	1.0359 (.262)
Individual 8	-.1328 ^b (.079)	-.0588 (.105)	1.0738 (.103)	1.0995 (.109)
Individual 9	.1157 ^c (.135)	.0262 (.178)	1.0457 (.206)	1.1199 (.230)
Individual 10	.3285 (.063)	.3558 (.075)	.0310 (.142)	.0197 (.144)
Individual 11	-.0273 ^d (.095)	-.0795 (.121)	1.1497 (.161)	1.1935 (.169)
Individual 12	.0225 ^a (.077)	.1036 (.098)	1.3400 (.143)	1.3512 (.149)
Individual 13	-.0705 ^a (.142)	.1010 (.228)	.7835 ^d (.167)	.8833 (.202)
Individual 14	-.1224 (.055)	-.1198 (.068)	1.0296 (.116)	1.0290 (.116)

¹ In this study, for all fourteen individuals, the range of the identification statistic $W\beta$ is between -4.9033 and .3154. $W\beta$ never being equal to, the structural model is identified for each individual.

^a Difference significant at $\alpha < .01$
^b Difference significant at $\alpha = .05$
^c Difference significant at $\alpha = .10$
^d Difference significant at $\alpha = .20$

for the first attribute, "size of the city". As shown by the results of the t-tests, the differences between the OLS estimates and the TSLS estimates of the halo effect β_j are statistically significant for most individuals. On the other hand, the estimates of the coefficient γ_j of B_{ij} do not change much. The results of the tests of significance of the differences between the two estimates for all three attributes are summarized in Table 5. This table shows that, in contrast to the Beck-

TABLE 5
Results of the tests of significance of the differences between the OLS estimates and TSLS estimates of β and γ for all three attributes

	$\alpha < .01$	$.01 < \alpha < .05$	$.05 < \alpha < .10$	$.10 < \alpha < .20$	Not significant	Total
β	14 (33.3%)	3 (7.1%)	3 (7.1%)	2 (4.8%)	20 (47.6%)	42 (100.0%)
γ	4 (9.5%)	1 (2.4%)	0 (0%)	3 (7.1%)	34 (81.0%)	42 (100.0%)

with and Lehmann results, our estimates of γ_j do not change much, but significant differences occur in the estimates of the halo effect. This means that the TSLS halo estimates cannot be accepted.

As to the attitude equation (1), Beckwith and Lehmann summarize their results by stating: "(T)he attitudes of less than half of the respondents could be explained with positive R^2 's by the TSLS estimated model. We concluded that the linear attribute model is false for at least a substantial fraction of the respondents in this study" (Beckwith and Lehmann, 1975). They argue that "the removal of the halo effect, as in the TSLS estimates, reduced the diagnostic and explanative power of (the attitude equation) for many people" (Beckwith and Lehmann, 1976). It should be mentioned at this point that very small values of the multiple correlation coefficient do not indicate a "poor fit" or lack of significance of the set of explanatory variables in the structural equation. As R^2 is constrained only to the interval $(-\infty, 1)$ large negative values can easily occur, but "negative values certainly do not discredit the model in any meaningful way" (Basmann, 1962; Tomek, 1970). Hence, the Beckwith and Lehmann conclusions seem to be questionable.

In order to assess the robustness of the Beckwith and Lehmann model of the effect of attitude on belief, the "average belief" variable was replaced by the "true" value for all three attributes, and the OLS estimates of the halo effect in the two alternative models were compared. The results which are not reported here show that both estimates are generally very close. On the other hand, the explanatory power of the Beckwith and Lehmann model is generally higher than that of the alternative model. Hence, the effect of the "average belief" variable B_{ij} is to provide an upward bias in the level of explanatory power of the belief equations. This may be partly due to the similar nature of the explanatory variable and the dependent variable.

In conclusion, a single equation model of the effect of attitude on beliefs about cities seems to be well specified according to our test results. Furthermore the

specification of the model - whether a single or a simultaneous equation - seriously affects the halo estimates. A similar type of model testing procedures could be carried out in another context, such as the modeling of the relationship between sales and advertising. In general, these testing methods should prove useful to the assessment of the functional relationships between marketing variables.

(February, 1976).

References

- R. L. Basmann, Letter to the Editor, Econometrica, 30 (October, 1962), 824-6.
- Neil E. Beckwith and Donald R. Lehmann, The Importance of Halo Effects in Multi-Attribute Attitude Models, Journal of Marketing Research, 12(August, 1975), 265-275.
- Neil E. Beckwith and Donald R. Lehmann, Halo Effects in Multi-Attribute Attitude Models: An Appraisal of Some Unresolved Issues, Journal of Marketing Research, 13 (November, 1976), 418-21.
- Joel Huber and William James, The Marginal Value of Physical Attributes: A Dollarmetric Approach, Presented at the Seventh American Marketing Association Attitude Conference, Hilton Head, South Carolina (February, 1976).
- Joel Huber and Bill James, A Theory of Halo, Working Paper, Krannert Graduate School of Management, Purdue University (March, 1976).
- Johnny K. Johansson, Douglas J. MacLachlan and Richard F. Yalch, Halo Effects in Multi-Attribute Attitude Models: Some Unresolved Issues, Journal of Marketing Research, 13(November, 1976) 414-7.
- E. A. Pessemier and R. D. Teach, A Single Subject Scaling Model Using Judged Distances Between Pairs of Stimuli, Paper No. 143, Institute for Research in the Behavioral, Economic and Management Sciences, Krannert Graduate School of Management, Purdue University, 1966.
- J. B. Ramsey, Tests for Specification Errors in Classical Least Squares Regression Analysis, Journal of the Royal Statistical Society, Ser. B, 31, 1969, 350-71.
- James B. Ramsey, Classical Model Selection Through Specification Error Tests, in Frontiers in Econometrics, ed. by Paul Zarembka (New York: Academic Press, 1974).
- J. B. Ramsey and R. Gilbert, A Monte Carlo Study of Some Small Sample Properties of Tests for Specification Error, Econometrics Working Paper, No. 6813, Michigan State University, East Lansing, Michigan, 1969.
- J. B. Ramsey and R. Gilbert, A Monte Carlo Study of Some Small Properties of Tests for Specification Error, Journal of the American Statistical Association, 67 (March 1972) 180-6.
- S. S. Shapiro and M. B. Wilk, An Analysis of Variance Test for Normality (Complete Samples), Biometrika, 52, 1965, 591-611.
- H. Theil, The Analysis of Disturbances in Regression Analysis, Journal of the American Statistical Association, 60(December, 1965) 1067-79.
- W. G. Tomek, R^2 in TSLs and GLS Estimation, American Journal of Agricultural Economics, 55(November, 1973), 670.
- Dick R. Wittink and Reiar Ness, Halo Effects in Multi-Attribute Attitude Models: A Caveat, Working Paper, Graduate School of Business, Stanford University

A STUDY OF THE AMOUNT OF HALO
IN THE PERCEPTIONS OF AUTOMOBILES

William L. Moore, Columbia University
William L. James, University of Alabama

Abstract

Only a small amount of halo was found in attribute judgments of automobiles using a methodology similar to Beckwith and Lehmann. The amount of halo found in perceptual spaces formed through either attribute or similarity judgments was found to be about equal and low in both cases.

Introduction

Multiattribute models have been used by marketers in an attempt to explain the preferences for objects by the amount of certain attributes that the objects are perceived to possess. While the studies have found that the models generally have good predictive power, a disturbing aspect of these studies has been that respondents have a tendency to rate objects they prefer higher than expected on desirable attributes.

This tendency generally increases the within subject correlation between belief scores on different attributes and it also increases the within subject correlation between the belief scores on attributes and the preference measure. This limits the multiattribute model's value as a diagnostic tool because the dimensionality of the attribute structure of the product class is confounded, as is the analysis of brand strengths and weaknesses. Similarly, if halo effects are present in data used to build perceptual spaces, they will be distorted and attempts to position products in "halo spaces" could be dangerous.

The purposes of the paper are twofold. First, is to study the amount of halo in attribute judgments of automobiles in a manner similar to Beckwith and Lehmann (1975) to determine if their findings are replicated in a very different product category. Second, to compare the amount of halo found in perceptual spaces built with similarity judgments with spaces constructed from attribute ratings.

Background

Psychologists have been aware of the presence of halo for over 50 years (Thorndike, 1920). Much of their work has focused on methods of measuring the amount of halo in trait ratings. This has been done by measuring the correlations between trait ratings (Symonds, 1925; Bingham, 1939; and Keaveny and McGann, 1975) and by measuring the variance of ratings across traits for a given object (Brown, 1968, and Wilkie, McGann, and Reibstein, 1973). A review of their studies is in Huber and James (1976) and won't be repeated.

Beckwith and Lehmann (1975) used a series of regression equations to study beliefs about the attitudes toward television shows. In the first equation they modeled an individual's attitude toward a T.V. program as a function of his beliefs about the program and the sample average attitude toward it

$$A_i = \sum_{j=1}^N \omega_j B_{ij} + \gamma A_i^* + u_i \quad (1)$$

where

A_i is the individual's attitude toward the i th program,

B_{ij} is the individual's belief about the i th program on the j th attribute,

A_i^* is the average attitude toward the i th program,

ω_j is the weight of the j th attribute,

γ is the weight of the average attitude, and

u_i is the disturbance term.

In the remaining equations the individual's beliefs about the attribute levels possessed by a program were modeled as functions of the sample average beliefs about the attribute levels and the individual's attitude toward the program.

$$B_{ij} = \beta_j A_i + \gamma_j B_{ij}^* + u_j \quad j = 1, \dots, N \quad (2)$$

where

B_{ij} , A_i , and u_j are as above,

B_{ij}^* is the average belief about the i th program on the j th attribute, β_j is the weight of individual attitude, and γ_j is the weight of the average belief.

The coefficients of the seven equations (one for attitude and one for each of the six attributes) were estimated using ordinary least squares. These equations were then estimated as a simultaneous system. While the simultaneous system was under identified for some individuals, similar parameter estimates were obtained with the two methods.

Beckwith and Lehmann found that the variation in an individual's beliefs about the attribute levels of a program could be modeled as a function of the sample average beliefs about the attribute levels and the individual's attitudes toward the program. While the regression coefficients associated with the average beliefs were usually larger than the coefficients associated with overall attitudes, they were of roughly the same size. These results indicated that a person's perception of or beliefs about an object can be explained almost as well by knowing what his attitude toward the object is as by knowing what the average perception of the object is.

There has been some criticism of the simultaneous equation methodology and behavioral interpretation (Johansson, MacLachlan, and Yalch, 1976). However, there seems to be little disagreement that there is a systematic relationship between average attribute beliefs, individual attitudes and individual attribute beliefs. This is not to imply that people know what the average beliefs are and that they consciously distort their judgments, but just that there is some systematic variation in individual attribute beliefs.

Data

Judgments on perceptions and preferences for ten automobiles were collected from forty junior and senior marketing students. Each student rated each automobile on ten attributes using a ten point Likert scale. Similarity judgments were also collected on a ten point scale. The ten automobiles and the ten descriptive attributes are given in Table 1.

TABLE 1
List of Brands and Attributes

Brands	Attributes
Cadillac	Gas mileage
Camaro	Style and appearance
Chevrolet Impala	Driver comfort
Ford Granada	Popularity with friends
Mercury Montego	Sensibleness for people on budgets
MG	Power and performance
Mustang	Dealer support
Toyota Corona	Value
Volvo 144	Handling
Vega	Purchase price

Preference measures were collected using Pessemier's dollar metric technique (Pessemier, Burger, Teach, and Tigert, 1971). Each student was asked which one of a pair of automobiles would be purchased upon graduation if he had to choose between those two. Then he was asked how much would the price of his choice have to increase before he would change his decision.

Subjects were first screened on their stated knowledge of the cars in the survey. Then they were screened based on a lack of pattern filling out the attribute ratings. Finally, these preference judgments were converted into the best fitting linear preference scale and only those subjects whose preference judgments were significantly different from random (Bechtel, 1967, Pessemier and Teach, 1970) were kept. Six subjects were eliminated through this screening.

Amount of Halo in Attribute Ratings

The first objective of this study is to determine if the results of Beckwith and Lehmann (1975) were idiosyncratic to the T.V. data, or if these findings are typical of the amount of the halo to be found in attribute judgments of other product classes. In doing this, the individual's beliefs about attribute levels of a brand were modeled as a linear function of the average beliefs about the attribute levels of the brand and the individual's preference for the brand. This is equation (2) above.

These regressions were estimated separately for each individual using OLS. The individual coefficients were aggregated across all the subjects and these aggregate results are presented in Table 2.

TABLE 2
Estimated Coefficients of the Belief Equations

Attribute	Average Belief Coefficients		
	$\bar{\gamma}_j$	$ t $	Number of times significant out of 34
Gas mileage	.855	4.31	34
Style & appearance	.648	2.84	24
Driver comfort	.723	3.66	25
Popular with friends	.441	1.77	12
Sensible for people on a budget	.470	1.77	15
Power & performance	.658	3.55	27
Dealer support	.530	2.01	19
Value for the money	.669	2.80	24
Handling	.654	2.40	17
Price	.633	2.28	21

Attribute	Overall Attitude Coefficients		
	$\bar{\beta}_j$	$ t $	Number of times significant out of 34
Gas mileage	.000	.94	4
Style & appearance	.055	1.23	6
Driver comfort	.027	1.09	6
Popular with friends	.028	.87	1
Sensible for people on a budget	-.105	.59	1
Power & performance	-.001	.79	2
Dealer support	.108	1.03	2
Value for the money	.086	.93	5
Handling	.034	.81	2
Price	.040	.85	4

The average standardized regression coefficients and average absolute values of the t statistics are much larger for the belief variable than the attitude variable for every attribute. In addition to the regression coefficients associated with the average belief measures being much larger, they are also significantly different from zero a much higher proportion of the time in the individual regressions than the coefficients associated with attitude.

In contrast to the Beckwith and Lehmann study, the halo effect appears to be relatively unimportant.

Halo in Product Spaces

If there is a halo effect biasing attribute judgments, then product spaces constructed using either discriminant analysis (Johnson, 1971) or factor analysis (Urban, 1975) would also contain halo. The effects of halo on judgments of overall object similarity has not been researched. However, it has been hypothesized (Beckwith and Lehmann, 1975) that product spaces built from similarity judgments would contain less halo than product spaces constructed using attribute judgments. This hypothesis was explored by constructing both aggregate and individual perceptual spaces then modeling the location of an object along a dimension in the individual space as a function of the object's location on the corresponding dimension in the aggregate space and the individual's attitude toward that object. This gave an equation analogous to equation (2) above.

$$X_{ij} = \beta_j A_i + \gamma_j X^*_{ij} + u_j \quad (3)$$

where

X_{ij} is the location of the i th object on the j th dimension of the individual space,

X_{ij}^* is the location of the i th object on the j th dimension of the aggregate space,

A_i is the individual's attitude toward the i th object,

U_j is the disturbance term, and

β_j and γ_j are regression coefficients.

Similarity Judgments

Similarity judgments for each of the 34 respondents were transformed into individual perceptual spaces using KYST (Kruskal, Young, and Seery, 1973). An aggregate similarity matrix was formed by averaging the pair-wise similarities over all the respondents. This matrix was used to form the aggregate perceptual space. Each of the individual spaces were rotated and reflected to a position of maximum congruence with the average space using a program based on Schönemann's solution to the Orthogonal Procrustes Problem (Schönemann, 1966). Then equation (3) was used to determine the relative importance one individual's preference and average perceptions in modeling individual perceptions.

Attribute Ratings

A reduced space description of the attribute ratings was formed through discriminant analysis. In this application of discriminant analysis, the brands under study were the groups and the discriminant functions were the linear combinations of the attributes that maximally separate the brands in the produce space. The location of the i th brand in the aggregate produce space X_i^* was found by multiplying the matrix, D , whose columns contain the discriminant vectors, times the vector of average attribute ratings B_i .

$$X_i^* = D' B_i \quad (4)$$

Similarly, the location of a brand in an individual product space, X_i , was found by multiplying the matrix D times the vector, B_i , containing that individual's beliefs about the attribute ratings for the brand

$$X_i = D' B_i \quad (5)$$

Equation (3) was used to determine the relative ability of individual preferences and average perceptions to model individual perceptions. These regressions were run for two dimensions in the product spaces for each individual and the results were aggregated across individuals and presented in Table 3.

TABLE 3
Estimated Coefficients of the
Product Space Coordinates

Model	Dimension	Average Coordinate Coefficients	
		$\bar{\gamma}_j$	$\bar{ t }$
Attribute	1	.864	6.47
Similarity	1	.824	5.71
Attribute	2	.743	3.81
Similarity	2	.492	1.81

Model	Dimension	Overall Attitude Coefficients	
		$\bar{\beta}_j$	$\bar{ t }$
Attribute	1	-.04	1.18
Similarity	1	-.102	1.39
Attribute	2	-.119	1.11
Similarity	2	.113	.99

In both methods of modeling perceptions, the average perceptions dominate the effect of preference. Comparing the t statistics, the average beliefs are a stronger influence in the attribute models and the effect of the preference component appears to be about the same in either case. It appears that there is little distortion due to halo in either method of constructing produce spaces and there would be little danger of positioning a product in a "halo space".

Conclusions

While these results aren't necessarily generalizable to the general public or to other product classes, this paper has shown that the amount of halo in attribute models may be quite low. These results also indicate that there may be only a very small danger of positioning some products in halo spaces.

Previous studies have indicated that certain types of attributes may be more susceptible to halo than other types: those without a clear physical analogue (Huber and James, 1976) vague, ambiguous and less important (Beckwith and Lehmann, 1975). Two more types of attributes that may give excessive halo are affective attributes (style and appearance) and the attributes that people have little knowledge of (dealer support).

It is also possible that some product classes generally have a greater degree of distortion due to halo than others. Frequently purchased goods may be routinely purchased without much thought of the attribute levels that different brands possess. When people are asked to make attribute judgments about these products, their responses may be more distorted by halo than their responses about the attribute levels of products that are purchased with more thought. Halo may be more of a problem in product classes where the differences between brands are fairly small, but much less of a problem in product classes with significant differences.

A useful area for future research is a thorough investigation of the amount of halo in attribute judgments of different produce classes. Also, while this study has shown that there is little difference in the amount of halo in attribute and similarity judgments, this topic deserves further study.

REFERENCES

- Gordan G. Bechtel, "The Analysis of Variance and Pair-wise Scaling," Psychometrika, Vol. 32(March 1967), pp. 47-65.
- N. E. Beckwith and D. L. Lehmann, "The Importance of Halo Effects in Multi-Attribute Attitude Models," Journal of Marketing Research, Vol. 12(August 1975), pp. 265-275.
- W. V. Bingham, "Halo, Invalid and Valid," Journal of Applied Psychology, Vol. 23(1939), pp. 221-228.
- Eva Brown, "Influence of Training, Method and Relationship on the Halo Effect," Journal of Applied Psychology, Vol. 52(July 1968), pp. 195-199.
- Joel C. Huber and William L. James, "A Theory of Halo," working paper Krannert Graduate School of Industrial Administration, Purdue University, March, 1976.
- W. L. James and F. S. Carter, "Halo Effects in Location Preferences," Paper presented at Albert Haring Symposium, April, 1976.
- J. K. Johansson, D. L. MacLachlan and R. F. Yalch, "Halo Effects in Multiattribute Models: Some Unresolved Issues," Journal of Marketing Research, Vol. 13(November, 1976), pp. 414-417.
- R. M. Johnson, "Market Segmentation: A Strategic Management Tool," Journal of Marketing Research, Vol. 8 (February 1971), pp. 13-18.
- Timothy Keaveny and Anthony McGann, "A Comparison of Behavioral Expectation Scales and Graphic Rating Scales," Journal of Applied Psychology, Vol. 60(December 1975), pp. 695-703.
- J. B. Kruskal, F. W. Young and J. B. Seery, "How to Use KYST, A Very Flexible Program To Do Multi-dimensional Scaling and Unfolding," unpublished manuscript, Bell Telephone Laboratories, 1973.
- E. A. Pessemier, P. C. Burger, R. D. Teach and D. J. Tigert, "Using Brand Preference Scales to Predict Brand Choice," Management Science, Vol. 6 (1971), pp. 371-385.
- E. A. Pessemier and R. D. Teach, "Disaggregation of Analysis of Variance for Paired Comparisons: An Application to a Marketing Experiment," Institute for Research in the Behavioral, Economic, and Management Sciences, Krannert Graduate School of Industrial Administration, Purdue University, Paper No. 282, August 1970.
- P. H. Schönemann, "A Generalized Solution of the Orthogonal Procrustes Problem," Psychometrika, Vol. 31(March 1966), pp. 1-9.
- Percival Symonds, "Notes on Ratings," Journal of Applied Psychology, Vol. 9(1925), pp. 188-195.
- E. L. Thorndike, "A Consistent Error in Psychological Ratings," Journal of Applied Psychology, Vol. 4(January 1920), pp. 25-29.
- G. L. Urban, "Perceptor: A Model for Product Design," Management Science, Vol. 21(April 1975), pp. 858-871.
- W. L. Wilkie, J. M. McCann and D. J. Reibstein, "Halo Effects in Brand Belief Measurement: Implications for Attitude Model Development," paper presented at the Annual Association for Consumer Research Conference, October 1973.
- W. L. Wilkie, and E. A. Pessemier, "Issues in Marketing's Use of Multi-Attribute Models," Journal of Marketing Research, Vol. 10 (November 1973), pp. 428-441.

EMPIRICAL EVIDENCE OF HALO EFFECTS IN
STORE IMAGE RESEARCH BY ESTIMATING TRUE LOCATIONS*

Neil E. Beckwith, University of Pennsylvania
U. Victor Kubilius (Ph.D. Candidate), University of Pennsylvania

Abstract

Individuals' judgments and responses often reflect many factors, such as their overall attitude toward the object being judged, its popularity, their familiarity with it, etc. These biases are similar to the well-known halo effect—an individual's tendency to bias his responses about an object on any specific attribute by his general, overall (global) impression. This article describes a procedure for estimating locations of the judged objects corrected for these types of halo-like effects.

Introduction

Marketing researchers and psychologists frequently attempt to obtain evaluations or locations of particular alternatives (individuals, objects, concepts, or brands) on particular attributes (traits, factors, variables, or characteristics), as in image studies, attitude models, and product attribute positioning models (Fishbein and Ajzen, 1975; Huber and James, 1977; Rosenberg, et al, 1960). Convenient estimators of these locations are the averages of many individuals' responses on the attribute. However, individuals may commingle their knowledge about the objects on many different attributes, and thereby provide response ratings on each attribute which are biased. Such biases may be due to their evaluations on other attributes, their overall evaluation of the object, its popularity, their usage or familiarity with it, their beliefs about other people's assessment of the object (peer attitude), cognitive dissonance (Festinger, 1957), or consistency maintenance (Rosenberg, et al, 1960). The resulting averages may thereby be biased.

The halo effect is well recognized by psychologists and has been defined as the "tendency in rating to be influenced by general impression or attitude when trying to judge separate traits" (English, 1934). Additional halo-like systematic response biases may be associated with other sources.

We demonstrate a method for obtaining estimates of the locations of objects on particular attributes. We refer to these as estimates of the true or actual location of the objects on the attribute, even though the attribute may be subjective.

Basically, our objective is to use individuals' responses, which indicate their rating of an object on an attribute, to construct an estimate of the unmeasured (true or actual) location of the object on the attribute. We assume that individuals' responses are a function of the (true) location as well as a function of other variables. First, we estimate each individual's response function and then use an iterative technique to estimate the (true) locations of the objects across all individuals.

The store image data which was conveniently available to

*The authors are indebted to Charles D. Goldfine and Karen E. Jakubowski (1977), who originally collected the data and made it available to us for this further analysis, and to Randy Batsell, who critically reviewed an earlier draft.

us has ordinal (ranking) rather than interval (rating) scales on some variables. Since the method we use assumes intervally scaled data, the empirical example is included only for illustration of the methodology. We draw no substantive inferences from it.

Background

The halo effect is well known to psychologists. Thorndike (1920) noted that an individual tends to rate another person on an attribute by his perception of the person on some other attributes. Similar halo-like effects occur in many other circumstances. Bass and Talarzyk (1972) reported that respondents indicated higher average belief scores on attributes for their preferred brands than for less favored brands. Beckwith and Lehmann (1975) demonstrated that respondents also halo their responses when asked to rate television shows as commonly done in the product positioning attribute models. Bass and Wilkie (1973) suggested that the halo effect on indicated brand beliefs may partially explain why importance weights do not seem to add significantly to the predictive performance of such attitude models. In contrast, Moore and James (1977) found halo effects to be relatively unimportant in students' ratings of automobiles. A summary of halo effect research and implications is available elsewhere (Beckwith, Kassarian, and Lehmann, 1977).

Store image studies provide a convenient example of attribute (i.e., trait) measurement problems and are of practical importance since objective measures on many of the attributes are difficult or impossible to obtain.¹

Two recent investigations provided considerable motivation for the present study. Huber and James (1976) pointed out that cross-sectional averages of halo-biased ratings on an attribute are also biased because the alternative objects (such as stores or brands) will generally be favored by different proportions of the respondents. Peterson (1976) found such a haloing effect in six studies of retail store images where average ratings on many attributes were correlated with the market shares of the retailers in their geographic marketing areas. We attempt to obtain less biased estimates of the objects' locations on attributes by considering both relative usage and overall attitudes in the sample of respondents.

Model

In this section we develop an example of a class of models useful for estimating the true (or actual) locations of objects on specific attributes. The example selected is both rather simple and readily extendable to more elaborate functional forms if warranted.

¹Examples of store image studies include those reported by Hawkins, Albaum, and Best (1975-76), Jain and Etgar (1976-77), Kunkel and Berry (1968), Lessig (1972), Linquist (1974-75), May (1971), Sharma and Doyle (1977), and Staples and Lockander (1975-76). Readers are also referred to the Special Store Image Issue of the Journal of Retailing (Winter 1974-75).

As a simple illustration we examine a model having only three components.² First, each object j has an actual location T_{jk} on each attribute k . T_{jk} may be thought of as the true location of the object on the attribute. Respondents may have been exposed to the object, and thereby exposed to cues of T_{jk} (see Castellan, 1973). The analyst is attempting to estimate T_{jk} .

Second, individual i 's evaluation E_{ijk} of object j on each attribute k reflects influences of many other variables besides T_{jk} . For purposes of this example, assume a simple linear transformation:

$$E_{ijk} = b_{i0k} + b_{i1k}A_{ij} + b_{i2k}U_{ij} + b_{i3k}T_{jk} \quad (1)$$

where A_{ij} = individual i 's overall evaluation of object j ,

U_{ij} = individual i 's usage of object j ,

$b_{i \cdot k}$ = individual i 's weighting coefficients,

T_{jk} = location of object j on attribute k .

The $b_{i1k}A_{ij}$ component corresponds to the usual halo effect. The $b_{i2k}U_{ij}$ component is included as an example of other halo-like biasing effects.³

Third, we assume that each individual's response B_{ijk} is a linear function of his evaluation E_{ijk} plus a disturbance e_{ijk} :

$$B_{ijk} = a_{i0k} + a_{i1k}E_{ijk} + e_{ijk} \quad (2)$$

Each individual uses his own linear scaling transformation with coefficients a_{i0k} and a_{i1k} . Disturbances e_{ijk} are independently distributed with mean zero and idiosyncratic variance σ_{ijk}^2 . Basically, this is a linear accommodation of the different response scales which individuals may employ to determine their own responses B_{ijk} based on their own evaluations E_{ijk} .

Substituting (1) into (2) yields the reduced form equation:⁴

$$B_{ijk} = \beta_{i0k} + \beta_{i1k}A_{ij} + \beta_{i2k}U_{ij} + \beta_{i3k}T_{jk} + e_{ijk} \quad (3)$$

The variables B_{ijk} , A_{ij} , and U_{ij} are respondent i 's indicated responses (data). However, T_{jk} , not empirically observable, is to be estimated. A naive estimator of T_{jk} is \bar{B}_{jk} , the average across all i respondents (Beckwith and Lehmann, 1975). However, Huber and James (1976) have pointed out that such an estimator is biased

²The more general linear model is presented at (4). Here we display the special case used as an example.

³In our empirical work we originally attempted to include three separate effects: Store of Last Purchase, Store of Most Frequent Purchase, and Most Convenient Store. However, for most subjects these variables were too collinear. We settled on the single usage measure, Store of Most Frequent Purchase. Similar results were obtained with all three measures of usage except for the Value attribute.

⁴Where $\beta_{i0k} = a_{i0k} + a_{i1k}b_{i0k}$, $\beta_{i1k} = a_{i1k}b_{i1k}$, etc.

because it neglects the differing fractions of respondents favoring each brand. Thus, we elect to estimate T_{jk} cross-sectionally.

Estimation Method

Consider a simple response function for individual i on one attribute k . Dropping the k subscripts for the moment:

$$B_{ij} = \sum_{\ell} \beta_{i\ell} X_{ij\ell} + \beta_{iT} T_j + e_{ij} \quad j = 1, \dots, J \quad (4)$$

where i 's response B_{ij} about object j is determined as a function of L independent intervally scaled variables $X_{ij\ell}$, $\ell = 0, \dots, L$ (including the intercept), and also as a function of the location T_j of object j . The independent variables $X_{ij\ell}$ could include the overall evaluation of the object j (like-dislike), preference, popularity of market share, peer attitudes or i 's belief about them, familiarity, etc. We seek to obtain estimates \hat{T}_j for each $j = 1, \dots, J$. These initial guesses could be chosen arbitrarily. We could use them to estimate all the β s in (4) for an individual i by an ordinary least squares (OLS) regression across the J objects, or perhaps by some other method. If we then examine i 's residuals $R_{ij} - \hat{R}_{ij}$ we would find them to be positive for some objects and negative for others. Assuming $\hat{\beta}_{iT} > 0$, we would then tend to guess that the corresponding original guess about T_j was too small (i.e., too negative) for those objects j with $\hat{e}_{ij} > 0$ and that the corresponding original guess about T_j was too large (i.e., too positive) for those objects j with $\hat{e}_{ij} < 0$.⁵

Instead of using \hat{e}_{ij} s from just one individual to improve the guesses about the T_j s, we suggest that the \hat{e}_{ij} s be obtained for many individuals and that a \hat{T}_j should be increased if $\sum_i \hat{e}_{ij} > 0$ and decreased if the sum of residuals for the object is negative, assuming the $\hat{\beta}_{iT}$ s are generally positive. The scale for \hat{T}_j and $\hat{\beta}_{iT}$ is arbitrary. To resolve this ambiguity, we simply constrain the \hat{T}_j scale to be standardized with zero mean, unit variance, and sign such that $\hat{\beta}_{iT} > 0$.⁶

Thus, this procedure starts from an original set of arbitrary values for the \hat{T}_j s. The β s of the response function are estimated for each individual by OLS. The \hat{T}_j s are then revised. We used the updating rule:

$$\hat{T}_j = \frac{\sum_i [B_{ij} - \sum_{\ell \neq T} \hat{\beta}_{i\ell} X_{ij\ell}]}{\sum_i \hat{\beta}_{iT}} \quad j = 1, \dots, J \quad (5)$$

The procedure is then repeated with the new \hat{T}_j s replacing the previous values until convergence, when $\sum_i \hat{e}_{ij} = 0$. Finally the signs of the \hat{T}_j s are selected such that $\hat{\beta}_{iT} \geq 0$, and the final \hat{T}_j s are scaled to have zero mean and unit variance.

In our example, the procedure converged to the same final locations regardless of the initial \hat{T}_j s arbitrarily selected. Since the procedure estimates (4) for each individual i separately, rather than conjoining them

⁵Of course, we could easily obtain all $\hat{e}_{ij} = 0$ $j = 1, \dots, J$ for any one particular individual i by rigging the \hat{T}_j s. However, we cannot usually find a set of \hat{T}_j s which will drive all $\hat{e}_{ij} = 0$ for many individuals.

⁶In retrospect we feel it might be better to constrain the estimated locations \hat{T}_j to a scale of zero-one, rather than standardizing to zero mean and unit variance. Negative locations seem to make communication of findings more difficult.

into some kind of a big pooled regression, each iteration of the estimation requires only a modest sized regression. Thus, the procedure can be performed on very modest computers, even for very large samples.

Data

The data used in this study were collected by Goldfine and Jakubowski (1977) as part of an advanced study project at The Wharton School, University of Pennsylvania. They investigated the homogeneity of store images and the degree of halo effect upon these store images within a convenience sample of professional women.

Although the methodology is designed to address substantive issues, our purpose here is only to display the methodological procedure. We use this convenient data base as a means of illustrating the application of the procedure. The reader is cautioned to consider this report as only illustrative of the types of results possible. The names of the stores (objects) are included only to facilitate understanding of the procedure.

The sample was obtained mainly on a convenience basis. It consisted of 43 women who resided in the greater Philadelphia area, and were employed in professional/administrative capacities or were studying toward that goal. Some of the respondents were acquaintances of the researchers. Responses from 9 of the women were not used due to missing data or errors. Also, individuals who indicated equal beliefs toward all 10 stores on an attribute were eliminated from analysis of that attribute since regression estimates of their weights were not possible. The 10 stores included diverse types familiar to many of the respondents.

Belief Responses B_{ijk}

The respondents were questioned about the 10 stores on 8 attributes.⁷ The attributes were selected on an a-priori basis. Thus, they are only illustrative of the types of attributes individuals might consider in making store evaluations. Each respondent i indicated her belief B_{ijk} about each store j on each attribute k using a five-point ("Outstanding" = 4, "Unacceptable" = 0) scale which we assumed to be an interval measure. The questionnaire items used to measure two of these attributes, Convenience and Fashionability of Merchandise, were worded such that the individual's differing ideal points (or home/work site) preclude assumption of the existence of a true location of a store on the attribute for all individuals.⁸ In addition, one of the usage measures, Most Convenient Store, also solicited convenience information. Consequently, estimates for these two attributes are not reported here.

Attitude A_{ij}

Each respondent indicated her overall evaluative ranking of each store ("Most Preferred" = 9, "Least Preferred" = 0, no ties). This measure is, of course, only ordinally scaled. However, we used it as if it were an intervally scaled variable for purposes of illustrating the methodology.

Usage U_{ij}

Three different measures of each individual's store

⁷See Table 2 for the list of stores and attributes. The standardized average beliefs \bar{B}_j across all individuals are displayed as \bar{B} in the tables. Raw, rather than standardized, belief, attitude, and usage variables were used in the estimation procedure.

⁸See the Appendix for questionnaire wording.

usage were available: (1) Store of Last Purchase, (2) Store of Most Frequent Purchase, and (3) Most Convenient Store. These variables were coded as dummies (an indicated store = 1, otherwise = 0). Since these three variables measure somewhat different constructs, we would advocate using all three within the analytic procedure. However, in this small set of data, the variables were too collinear to permit their simultaneous inclusion. For purposes of comparison, we included them each one-by-one as a single measure of usage in separate analyses. These usage variables do not meet the interval scale assumption.⁹

While these particular data from a convenience sample are not exactly consistent with the model's assumptions, they are adequate to simply illustrate the methodological procedure.

Illustrative Application

The Quality of Merchandise attribute is used in an example of the procedure. We selected the second operationalization (Store of Most Frequent Purchase) as the measure of usage. Other attributes and the importance of the halo effect are then examined in subsequent sections.

Quality of Merchandise Attribute

In lieu of the yet unavailable locations T_j we selected the cross-sectional averages \bar{B}_j for the initial arbitrary starting point. We then used OLS to estimate $\beta_{i,ks}$ in (3) for each particular individual (one at a time). The 10 stores were used as 10 observations. The OLS estimated coefficients (and standard errors) for one particular individual were:

$$B_{ijk} = -.48 - .03A_{ij} + .28U_{ij} + 1.48\bar{B}_{jk} + \hat{e}_{ijk},$$

(.18) (.07) (.36) (.19)

$$R^2 = .98 \text{ (unadjusted).}$$

This procedure is replicated for each individual in the sample. The resulting residuals do not sum to zero across individuals for each store. For example, here the sum of residuals $\sum_i \hat{e}_{ijk} = .10$ for the first store, Bonwit Teller. Using an updating rule, such as (5), the estimates of T_j s are revised away from \bar{B}_{jk} , and the OLS regressions are then repeated for all individuals. Eventually, the procedure converges to final estimates of the true locations \hat{T}_j . For this same individual the final estimated coefficients (and standard errors) were:

$$B_{ijk} = 1.64 + .15A_{ij} - .28U_{ij} + .92\hat{T}_{jk} + \hat{e}_{ijk},$$

(.23) (.05) (.34) (.13)

$$R^2 = .98 \text{ (unadjusted).}$$

After convergence the sum across individuals of the final residuals is zero for each store, e.g., $\sum_i \hat{e}_{ijk} = 0$, $j = 1, \dots, J$.

Although the analysis yields estimates like these for each individual, such results are too bulky to report in detail.¹⁰ Figure 1 displays the distribution of $\hat{\beta}$ s, standard errors, and t-statistics ($H_0: \beta_{ilk} = 0$) obtained for all 34 individuals. With only 10 objects

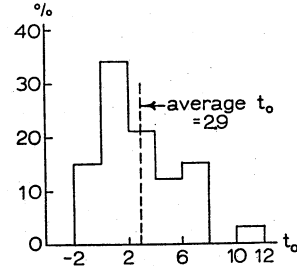
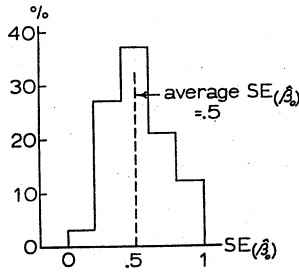
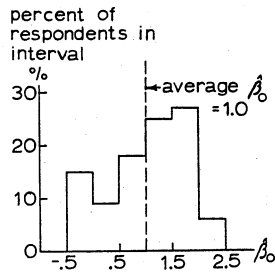
⁹For an individual having usage variable value of 1 for only one store, that store is essentially ignored in estimating the individuals' coefficients $\beta_0, \beta_1, \beta_3$, and determines $\hat{\beta}_2$. Some individuals indicated multiple stores.

¹⁰Final R^2 s exceeded initial R^2 s for about half of the respondents on all attributes.

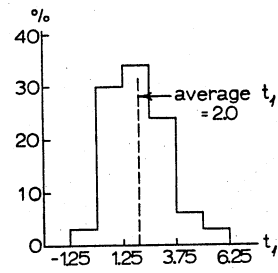
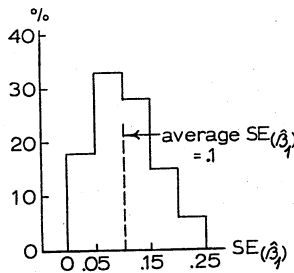
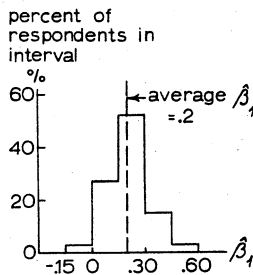
FIGURE 1

Distribution of Estimated Coefficients, Standard Errors, and t-Statistics for Explaining Individuals' Responses on the Quality of Merchandise Attribute^a

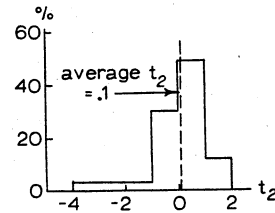
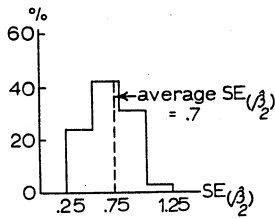
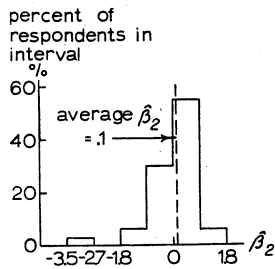
INTERCEPT COEFFICIENT, $\hat{\beta}_0$



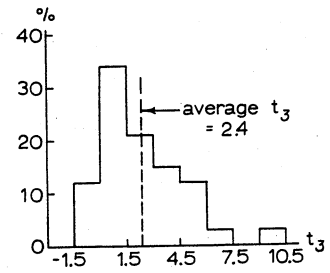
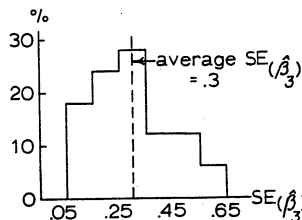
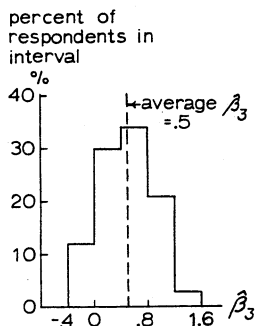
COEFFICIENT OF OVERALL ATTITUDE, $\hat{\beta}_1$



COEFFICIENT OF STORE USAGE, $\hat{\beta}_2$



COEFFICIENT OF TRUE LOCATIONS, $\hat{\beta}_3$



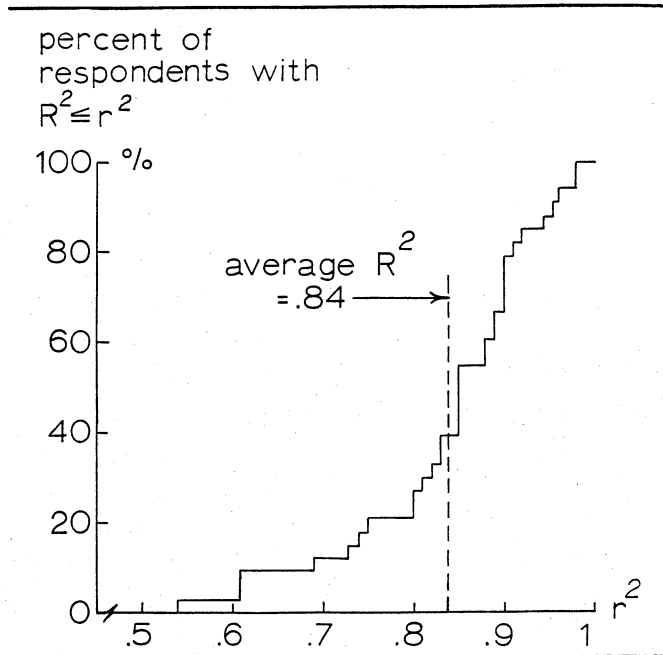
^aUsing the second usage measure (Store of Most Frequent Purchase).

(stores) the standard errors tend to be rather large. Even so, of the 34 t-statistics for $\hat{\beta}_1$, for example, 12 exceeded the critical value 2.447 for 6 degrees of freedom at the 95 percent confidence level. Since it is unlikely that this many estimated coefficients would have $|t| \geq 2.447$ by chance alone, we conclude that β_1 cannot be equal to zero for all individuals, although it could be zero for some individuals. Thus, we conclude that the halo effect cannot be ignored here.

Figure 2 displays the distribution of unadjusted R^2 obtained for these individuals, which ranged from .54 to .98. Averages of these statistics are summarized in the

FIGURE 2

Cumulative Distribution of R^2 Statistics for Explaining Individuals' Responses on the Quality of Merchandise Attribute^a



^aUsing the second usage measure (Store of Most Frequent Purchase), R^2 s are not adjusted.

bottom portion of Table 1, column 2. Columns 1 and 3 display similar average results for analyses using the other two measures of store usage: Store of Last Purchase (#1) and Most Convenient Store (#3), in place of Store of Most Frequent Purchase (#2).

The average response \bar{B}_{jk} and estimated true locations \hat{T}_{jk} of the 10 stores on the Quality of Merchandise attribute are displayed at the top of Table 1. Note that the estimated locations \hat{T}_{jk} were very similar to the average responses \bar{B}_{jk} for several of the stores. However, for Nan Duskin, Gimbels, and Wanamaker, the estimated locations \hat{T}_j were noticeably different from the average responses \bar{B}_j . Nan Duskin is one of the less frequented stores; none of the respondents mentioned it as being either the Store of Last Purchase or Store of Most Frequent Purchase. Thus, it is not surprising that the estimated \hat{T}_j is higher than \bar{B}_j for this store. Wanamaker was mentioned in this sample as one of the more frequently shopped stores (although less than Loehmann's). Thus, it is not surprising that the estimated \hat{T}_j is less than \bar{B}_j for this store. Both of these differences are consistent with the notion that average responses \bar{B}_j may be favorably biased estimates of T_j for stores (or brands) having larger market shares.

TABLE 1

Comparison of Results for Quality of Merchandise Attribute^a

STORE	AVERAGE RESPONSE \bar{B}	ESTIMATED LOCATIONS \hat{T}		
		1	2	3
Bonwit Teller	1.5	1.4	1.5	1.5
Casual Corner	-.1	.1	.1	.1
Clover	-1.6	-1.5	-1.5	-1.4
Gimbels	-.3	-.5	-.5	-.7
Lit Brothers	-1.1	-1.0	-1.0	-.9
Loehmann's	.4	.4	.4	.4
Nan Duskin	1.2	1.5	1.6	1.6
J. C. Penney	-1.2	-1.2	-1.2	-1.2
Strawbridge	.4	.3	.3	.2
Wanamaker	.8	.5	.3	.4
Estimated Intercept	$\bar{\beta}_0$	1.0	1.0	.9
	S.E.	.5	.5	.5
Estimated Coefficient of Overall Attitude	$\bar{\beta}_1$.2	.2	.2
	S.E.	.1	.1	.1
Estimated Coefficient of Usage	$\bar{\beta}_2$.0	.1	-.1
	S.E.	.7	.7	.6
Estimated Coefficient of Location	$\bar{\beta}_3$.5	.5	.4
	S.E.	.3	.3	.3
Average R^2 (not adjusted)		.8	.8	.8

^aThe average responses \bar{B} and the estimated locations \hat{T} are standardized across stores to zero mean and unit variance. The estimated location columns headed 1, 2, and 3 correspond to the three different measures of store usage: 1 = Store of Last Purchase, 2 = Store of Most Frequent Purchase, and 3 = Most Convenient Store.

Value Attribute

The same analysis was performed on the Value attribute generated by the questionnaire item: "Price should be considered in terms of value for money spent." The average responses \bar{B}_j on this Value attribute varied widely among stores. However, the estimated locations \hat{T}_j were roughly similar for all stores except Loehmann's (see Figure 3).

Loehmann's is rather distinct from the other stores. For example, it sells many items with labels excised. Loehmann's had the lowest estimated locations \hat{T} of any store on several other attributes: Quality of Store Personnel, Customer Services, and Store Atmosphere (see Table 2). Readers familiar with Loehmann's may find these estimates plausible, as do the authors.

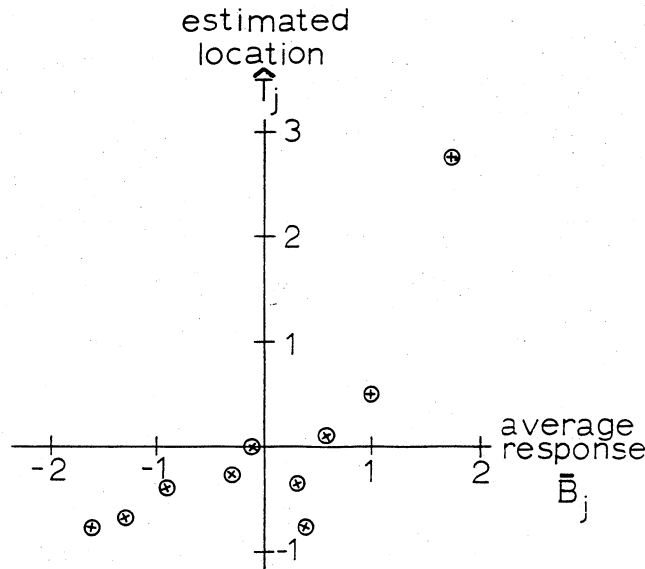
For many respondents the second measure of usage, Store of Most Frequent Purchase, was highly correlated with \hat{T}_j . This collinearity caused large standard errors for $\hat{\beta}_2$ and $\hat{\beta}_3$ for many individuals. Consequently, for this attribute we used the first measure of usage, Store of Last Purchase.

Other Attributes

Corresponding summary results for the other attributes are displayed in Table 2. For two of the attributes, Quality of Store Personnel and Customer Services, the estimated locations \hat{T} differ quite considerably from the average responses \bar{B} for many of the stores. The procedure converged in between 19 and 58 iterations for all the attributes except Value, which needed 162 itera-

FIGURE 3

Average Response \bar{B}_j and Estimated Locations \hat{T}_j for the 10 Stores on the Value Attribute^a



^aThe average responses \bar{B}_j and the estimated locations \hat{T}_j are standardized across stores to zero mean and unit variance. The average responses for the stores varied between 1.05 and 2.94 on the questionnaire Value scale before standardization. The first measure of usage, Store of Last Purchase, was used.

TABLE 2

Summary of Results for All Attributes^a

STORE	ATTRIBUTE												
	Value		Quality of Merchandise		Appeal of Advertising		Quality of Store Person		Customer Services		Store Atmosphere		
	\bar{B}	\hat{T}	\bar{B}	\hat{T}	\bar{B}	\hat{T}	\bar{B}	\hat{T}	\bar{B}	\hat{T}	\bar{B}	\hat{T}	
Bonwit Teller	.4	-.8	1.5	1.5	1.8	1.9	1.6	.8	1.6	1.2	1.7	1.7	
Casual Corner	-.3	-.3	-.1	.1	-.5	-.7	-.3	.0	-.4	-.5	-.1	-.0	
Clover	-1.6	-.8	-1.6	-1.5	-1.3	-.9	-1.6	-.3	-1.5	-.9	-1.3	-1.0	
Gimbels	-.1	.0	-.3	-.5	-.0	-.1	.2	.4	.1	.2	-.2	-.2	
Lit Brothers	-1.3	-.7	-1.1	-1.0	-1.3	-1.3	-.5	.9	-.7	.0	-1.0	-.9	
Loehmann's	1.8	2.8	.4	.4	-.3	.0	-1.1	-2.8	-1.5	-2.4	-.9	-1.3	
Nan Duskin	.3	-.4	1.2	1.6	1.3	1.7	1.0	.5	1.0	.8	1.5	1.8	
J. C. Penney	-.9	-.3	-1.2	-1.2	-.8	-.6	-.9	.0	-.4	.4	-.8	-.6	
Strawbridge	.6	.1	.4	.3	.3	-.2	.8	.3	.7	.5	.4	.1	
Wanamaker	1.0	.5	.8	.3	.8	.2	1.0	.3	.9	.7	.7	.3	
Estimated Intercept	$\bar{\beta}_0$		1.2		1.0		.9		.9		1.2		1.3
	S.E.		.4		.5		.6		.4		.4		.5
Estimated Coefficient of Overall Attitude	$\bar{\beta}_1$.2		.2		.2		.2		.2		.2
	S.E.		.1		.1		.1		.1		.1		.1
Estimated Coefficient of Usage	$\bar{\beta}_2$		-.1		.1		-.4		.2		-.2		.0
	S.E.		1.0		.7		.8		1.4		1.2		.8
Estimated Coefficient of Location	$\bar{\beta}_3$.4		.5		.5		.3		.4		.7
	S.E.		.3		.3		.3		.4		.4		.3
Average R ² (not adjusted)			.7		.8		.8		.6		.7		.8

^aThe average responses \bar{B} and the estimated locations \hat{T} are standardized across stores to zero mean and unit variance. For each attribute except Value, the results were somewhat similar, but not identical, for all three usage measures. Thus, results are only reported for the second usage measure (Store of Most Frequent Purchase) for all attributes except Value. The first usage measure (Store of Last Purchase) was used for Value because of collinearity between \hat{T} and usage measure #2.

tions. The average was 39 iterations before termination (when none of the T_j s changed more than 10^{-7}).

Importance of Halo Effect

In the previous section we concluded that the halo effect cannot be ignored here. The relative importance of the halo effect is indicated by comparing $avg_i(\sigma_{iA} \hat{\beta}_{i1})$, $avg_i(\sigma_{iU} \hat{\beta}_{i2})$, and $avg_i(\sigma_{iT} \hat{\beta}_{i3}) = \hat{\beta}_T$ for each of the attributes (see Table 3).

TABLE 3

Relative Importance of Halo Effect

Attribute	$\sigma_{iA} \hat{\beta}_{i1}$ Attitude (halo)	$\sigma_{iU} \hat{\beta}_{i2}$ Usage	$\sigma_{iT} \hat{\beta}_{i3}$ Location
Quality of Merchandise	.6	.02	.5
Customer Services	.5	-.05	.4
Appeal of Advertising	.6	-.13	.5
Store Atmosphere	.4	.01	.7
Quality of Store Personnel	.5	.07	.3
Value	.5	-.03	.4

The halo (overall attitude) and (true) location components are appreciable and roughly equal for all attributes. The usage component is negligible for all the attributes, evidently because of the low variances σ_{iU}^2 of the 0-1 measure of usage. Comparison between attributes is not possible here because the dependent variable, belief response, was not standardized.¹¹

Caveats

In retrospect, we are now concerned with several issues which were not adequately resolved by this study:

1. Simultaneity—We assumed that none of the right hand variables are determined as a function of the left hand variable. Our estimation is biased if this assumption is false. This type of analysis should be extended to allow for at least the possibility of simultaneous co-determination of overall attitude and the beliefs, perhaps along the lines of Beckwith and Lehmann (1975). Benmaor and Huber (1977) have argued that this specification error may be relatively tolerable, but we have not verified it for this example.
2. Intervals scaled data—We used available rank order data as if it were intervally scaled in order to simply demonstrate the methodology. Instead, intervally scaled data should be used. In addition, the convenience sample of respondents precludes making any substantive inferences from this particular data. However, these data are useful for demonstrating the methodology and illustrating the types of results which might be expected.
3. Individual coefficient nonnegativity—On a-priori grounds, individuals' coefficients $\hat{\beta}_{iTk}$ and $\hat{\beta}_{iAk}$ could be assumed to be nonnegative. By using OLS to estimate each individual's coefficients, we al-

¹¹In retrospect, we should have standardized the data to allow comparisons across attributes.

lowed these estimates to be negative, and many were. A linear programming-based individual estimation procedure might be used instead to obtain estimates constrained to being nonnegative, along the lines of Pekelman and Sen (1974) and Shocker and Srinivasan (1974). This might also reduce the collinearity problem.

4. Statistical properties—We have not demonstrated that the suggested procedure yields estimators T_{jk} which are necessarily no more biased than B_{jk} as estimators of T_{jk} . Presently we only point out that the procedure appears to yield plausible estimates in our example case, and that it seems reasonable to us that the bias will usually be reduced since the procedure explicitly includes effects which would otherwise be "omitted variables," and therefore generally biasing. Also, the OLS single equation standard errors and R^2 s reported for individuals are suspect since T_j is estimated along with the β coefficients.
5. Objective attribute recovery—We have not yet checked the ability of this procedure to retrieve accurate location estimates for attributes with objective physical measures.
6. Computational efficiency—Other computational schema may be found which estimate the true locations more economically. We have not yet attempted to reduce the computational burden of the procedure.
7. Underlying processes—The model, (1) and (2), used for this illustrative example seems to be an oversimplification of the individual's cognitive processes. We hope that other investigators will find this type of analysis useful in developing more complete understanding of such processes.

Summary

A methodology for estimating the location of objects on attributes or traits using individuals' biased response ratings was demonstrated. A procedure like this should be useful for assessing images of objects, such as stores, in image studies, or brand locations in product attribute analyses. The iterative procedure does require a substantial, although not prohibitive, amount of computation. The procedure was demonstrated on an example set of data for purposes of illustrating the methodology and the types of results which might be obtained. However, because the conveniently available data are not completely adequate, we caution against drawing substantive inferences from the example results. They are displayed solely to facilitate description of the methodology. We hope that this work will encourage other investigators to pursue more refined estimators of the unobservable locations or attribute levels of objects.

Appendix

Portions of the questionnaire are reproduced so that readers may judge for themselves the nature of the attributes measured.

Overall Attitude (A_{ij})

We are interested in your overall opinions of the following stores in a very general way. Please rank the stores numerically from 1 to 10 with 1 = most favored, 2 = second most favored, etc., with 10 meaning least favored. [The coding order was subsequently reversed, 9 to 0, for the analysis reported here.]

Store Usage (U_{ij})

At which of these stores did you most recently purchase an article of clothing for yourself? (If none, please specify) [Denoted Usage Variable #1.]

At which of these stores do you most frequently purchase clothing for yourself? (If none, please specify) [Denoted Usage Variable #2.]

Which of these stores do you consider most convenient (i.e., easiest to get to) to either your place of business or your home? (Include branch stores when applicable. If more, please specify) [Denoted Usage Variable #3.]

Belief Responses (B_{ijk})

We are interested in determining your opinions on a number of characteristics for several stores. Please indicate your opinion of each store on each attribute by checking the appropriate space. A short description of these characteristics precedes each section.

A. Quality of Merchandise—This refers only to the materials used in the manufacture of the garment and the garment's construction without any reference to price, fashionability, or any other factors.

B. Price—This should be considered in terms of value for money spent. [Price is referred to as the value attribute in the report.]

C. Quality of Store Personnel—Under this heading you should consider such factors as helpfulness, friendliness, promptness, courtesy, and any other interpersonal dynamic that is important to you.

D. Store Atmosphere—Store atmosphere includes the store's cleanliness, lighting, interior design, background music (if any), and floor plan.

E. Appeal of Advertising—In answering this question, you might consider whether or not the store's ads attract your eye in a newspaper or magazine as well as whether the ads are likely to result in your visiting the store in question.

F. Convenience of Location—(Consider branch stores when appropriate) In this category consider such factors as traffic, parking, proximity to public transportation, how frequently you find yourself in the area where the store is located, and so forth.

G. Customer Services—Customer services include credit availability and terms, return or exchange policies, alterations, rest room facilities, in-store restaurants, or any other particular service that you consider important.

H. Fashionability of Merchandise—In determining the fashionability of a store's merchandise, ask yourself whether the store carries merchandise that you like, not whether the store's merchandise could be featured in a fashion magazine.

Five response categories were provided for each store on each of the above beliefs: "Outstanding," "Highly Acceptable," "Acceptable," "Somewhat Acceptable," and "Unacceptable." These were subsequently coded 4 to 0 respectively.

References

A. H. Barton and R. W. Parsons, "Measuring Belief System Structure," Public Opinion Quarterly, 41 (Summer 1977), 159-180.

Frank M. Bass and W. Wayne Talarzyk, "An Attitude Model for the Study of Brand Preference," Journal of Marketing Research, 9 (February 1972), 93-96.

Frank M. Bass and William L. Wilkie, "A Comparative Analysis of Attitudinal Predictions of Brand Preference," Journal of Marketing Research, 10 (August 1973), 262-269.

Neil E. Beckwith and Donald R. Lehmann, "The Importance of Halo Effects in Multi-Attribute Attitude Models," Journal of Marketing Research, 12 (August 1976), 265-275.

Neil E. Beckwith, Harold H. Kassarian, and Donald R. Lehmann, "Halo Effects in Marketing Research: Review and Prognosis," in Advances in Consumer Research, 5 (1977 Proceedings).

Albert C. Bemmar and Joel C. Huber, "Econometric Estimation of Halo Effects: Single vs. Simultaneous Equation Models," in Advances in Consumer Research, 5 (1977 Proceedings).

N. John Castellan, Jr., "Comments on the 'Lens Model' Equation and the Analysis of Multiple-Cue Judgment Tasks," Psychometrika, 38 (March 1973), 87-100.

Horace B. English, A Student's Dictionary of Psychological Terms, fourth edition (New York: Harper Brothers, Inc., 1934).

Martin Fishbein and Icek Ajzen, Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research (Reading, Mass.: Addison-Wesley, 1975).

Charles D. Goldfine and Karen E. Jakubowski, "Store Image, Shopping Behavior, and the Halo Effect," unpublished paper, 1977.

Leon Festinger, A Theory of Cognitive Dissonance (Evanston, Ill.: Row-Peterson, 1957).

Del I. Hawkins, Gerald Albaum, and Roger Best, "Reliability of Retail Store Images as Measured by the Stapel Scale," Journal of Retailing, 52 (Winter 1975-1976), 31-38, 92-93.

Joel C. Huber and William L. James, "A Theory of Halo," working paper, Krannert School of Industrial Administration, Purdue University, March 1976.

_____, "A Concept of Halo," in Advances in Consumer Behavior, 5 (1977 Proceedings).

Arun K. Jain and Michael Etgar, "Measuring Store Image Through Multidimensional Scaling of Free Response Data," Journal of Retailing, 52 (Winter 1976-1977), 61-70, 95-96.

_____, Journal of Retailing, Special Store Image Issue, 50 (Winter 1974-1975).

John H. Kunkel and Leonard L. Berry, "A Behavioral Conception of Retail Image," Journal of Retailing, 32 (October 1968), 21-27.

P. F. Lazarsfeld and N. W. Henry, Latent Structure Analysis (Boston: Houghton-Mifflin, 1968).

- V. Parker Lessig, "Relating Multivariate Measures of Store Loyalty and Store Image," in Proceedings, American Marketing Association, 1972.
- Jay D. Lindquist, "Meaning of Image," Journal of Retailing, 50 (Winter 1974-1975), 29-38, 116.
- Eleanor G. May, "Image Evaluation of a Department Store," Marketing Science Institute, Cambridge, 1971.
- William L. Moore and William L. James, "A Study of the Amount of Halo in the Perception of Automobiles," in Advances in Consumer Research, 5 (1977 Proceedings).
- Dov Pekelman and Subrata K. Sen, "Mathematical Programming Models for the Determination of Attribute Weights," Management Science, 20 (April 1974), 1217-1229.
- Gene H. Peterson, "Design of Image Research," unpublished working paper, Deluxe Check Printers, Inc., St. Paul, January 1976.
- M. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson, and J. W. Brehm, Attitude Organization and Change, An Analysis of Consistency Among Attitude Components (New Haven: Yale University, 1960).
- L. G. Schiffman, J. F. Dash, and W. R. Dillon, "The Contribution of Store-Image Characteristics to Store-Type Choice," Journal of Retailing, 53 (Summer 1977), 3-14, 46.
- Alok Sharma and Peter Doyle, "A Model for Strategic Positioning in Retailing," in Barnett A. Greenberg and Danny N. Bellenger, eds., Contemporary Marketing Thought (Chicago: American Marketing Association, 1977), 10-14.
- Allan P. Schocker and V. Srinivasan, "A Consumer-Based Methodology for the Identification of New Product Ideas," Management Science, 20 (February 1974), 921-937.
- William A. Staples and William B. Locander, "Behaviorally Anchored Scales: A New Tool for Retail Management Evaluation and Control," Journal of Retailing, 52 (Winter 1975-1976), 39-48, 94-96.
- Edward L. Thorndike, "A Consistent Error in Psychological Ratings," Journal of Applied Psychology, 4 (January 1920), 25-29.

IDENTIFYING AND DEFINING CONSUMER NEEDS USING HUMAN FACTORS
AND MARKET RESEARCH TECHNIQUES

Thomas Cannon, Design Factors
Dr. Ronald W. Hasty, Colorado State University

Abstract

Consumer need identification and definition are areas whose importance in successfully defining new product concepts, evaluating existing products, and evaluating promotional materials effectiveness has not been fully appreciated by many in consumer research. This paper discusses the significance of these areas in reducing risk and uncertainty in new product development. The application of specific human factors and market research techniques to need identification and definition and the development of specific products using such techniques is discussed.

New Product Development Today

A recent paper (Crawford, 1977) raised the question "Why has the rate of new product success not climbed as a result of the many advances in marketing research technology over the past 25 years?" Crawford cites numerous studies which report failure rates up to 90%. The reasons for such failures range from poor timing, as in the case of the Edsel, to poor planning. However, the most predominant reason for failure, cited by all the studies examined, was "the lack of meaningful superior product uniqueness."

Crawford concludes that while market research techniques may be able to determine what consumers think of products once they are in the market, they are not particularly valuable in the development of truly new products or determining if people will buy the new products before they are introduced in the market.

The authors of this paper take the position that much of the failure to improve the success ratio of new products is traceable to gaps in the conventional product planning process. The major gap results from a lack of techniques suited to systematically identify and define consumer needs. The techniques necessary for these tasks must come from two sources. Those which are useful in consumer need identification come primarily from the field of market research and those suitable for consumer need definition come from the field of human factors research.

Figure 1 shows a typical new product planning process. Stage one is the development of objectives. Stage two is new product idea generation. Stage three screens the ideas for acceptability measured by the objectives set in stage one. Acceptable ideas undergo engineering development in stage four followed by market testing in stage five and commercialization in stage six.

In the process outlined, new product planning is basically a marketing activity. However, engineering and other non-marketing sources normally have responsibility for the generation and development of new product ideas. It is generally accepted in the marketing literature that marketing is responsible for developing idea sources, not ideas, and then coordinating the process required for screening them. (Shocker, 1969 & Larson, 1963)

Planning Process Gaps

The generation of new product ideas is an activity which suffers from a belief that all really successful new

products originate in a flash of creative insight. While such instances are not unknown it is evident that dependence on individual strokes of genius is not an adequate method of meeting the continuing corporate demand for new products. Most knowledgeable individuals would agree with this, yet brainstorming and other methods which attempt to increase the number of such creative flashes, are widely used. Such techniques continue to promote the myth that locked inside every individual is a "bright idea" for a product which would have wide appeal.

What is missing in such approaches is the recognition that any successful new product must be a solution to a consumer need. Once this is understood, it is apparent that the generation of new product ideas should be preceded by (1) the identification and (2) the definition of a need. Logically, only after this is accomplished can solutions to the need be developed through the exercise of creative abilities.

To solve the problem of filling this gap between stages one and two, techniques are needed for developing a specification of those characteristics required to provide consumer satisfaction in accomplishing a given end or task within a specific environment. Four factors must be addressed: (1) the degree of need, (2) characteristics of the consumer, (3) characteristics of the task, and (4) characteristics of the environment.

The degree of need can be defined as the consumers' perception that a discrepancy exists between a condition as it is and as it might be (Smith, 1973). The baseline for determining the degree of such discrepancy is the consumer's present condition. While attempts have been undertaken to make such measurements through the use of motivational research, this type of basic research has been generally neglected by market researchers. This is so, in part, because such measurements are difficult to make without specific analytical tools typically unused by the market researcher.

Utilizing a need characterization methodology to guide the generation of new product ideas can reduce the number of ideas to be screened in stage three to those which offer the greatest potential for satisfying Crawford's requirement of meaningful superior product uniqueness with particular emphasis on "meaningful." The purpose of stage three might be better understood if it were referred to as the screening of need solution concepts. This would keep the focus on an idea's usefulness to consumers and prevent the common mistake of concentrating upon the uniqueness aspect. Most product planning processes use stage three to evaluate concepts to determine if they have the potential to achieve profitability objectives, will result in greater market penetration, and achieve other business management goals. Emphasis on such measures can create an atmosphere in which evaluation of a concepts ability to truly meet the consumers defined needs is neglected.

Market researchers have frequently attempted to attack this problem by using concept testing. This technique attempts to determine if the potential user understands the idea of the proposed product, reacts favorably to it

and feels it answers their need (Luck, 1970). Concept testing, as presently practiced, has not proven to be very reliable in predicting the ultimate success of products (Tauber, 1975). One difficulty with most concept testing is that the consumer is not in a position where evaluations can be made in an educated manner. The consumer has no opportunity to experience the product in the actual use environment, but must make a subjective evaluation based on a verbal or visual representation of the concept. A second serious problem is that all product attributes which are important to the consumer may not be revealed during the concept test because the consumer is not able to predict his or her own behavior in relation to the product over time.

At present, decisions made in stage three are critical because they commit the firm to development in a particular direction thereby excluding alternative courses of action. And this is usually done on the basis of intuition, consensus, and judgment influenced by past experience.

What is missing is a method of allowing potential consumers to evaluate solution concepts and their characteristics in a realistic use environment before the decision to start product development activities is made. Only if this is done can the time consuming and expensive practice of developing products which are not marketable be eliminated. It is a fact that engineering departments are prone to expending resources on concepts which address problems which are not primary to users or in configurations which are not acceptable to users. In addition, the desire to demonstrate technological feasibility often results in solutions which fatally compromise consumer requirements.

Identifying Need With Market Research Techniques

Many existing market research techniques can be readily adapted to identifying various market segments and ranking the segments in terms of profit potential, suitability for existing distribution channels, etc. Identifying specific needs of these market segments can also be accomplished using existing techniques. However, because in the past these techniques have been used almost exclusively for evaluating existing or proposed product attributes, market researchers must reorient their perspective to one which does not focus on the product and its attributes but rather on basic consumer need structures. Research techniques such as depth interviewing, projective methods of word association, thematic application tests, sentence completion, and so forth have provided ways in which information can be obtained from consumers concerning their thoughts and feelings about product characteristics. The task now facing market research is to adapt these proven techniques to provide specific information concerning problems perceived by consumers in their present methods of achieving goals. The authors have utilized such tried and true techniques as in-depth interviews, questionnaires, and focus groups to identify specific areas of need. These and other techniques are also useful in defining with precision the psychological or "felt" needs of consumers.

Defining Physical Needs With Human Factors Research Techniques

Techniques from the field of human factors research provide a complementary capability which allows researchers to answer questions about physical characteristics which are critical to consumer satisfaction. These questions have to do with the range of physical capabilities exhibited by the market segment under consideration, the ways in which consumers presently perform various tasks and the role which environmental factors play in

the accomplishment of tasks. While there are a number of tested human factors techniques which can be utilized to define consumer need, those which are most useful come under the general category of observational methods. In addition to personal observation of consumers the human factors specialist utilizes slow motion and time lapse photography, video recording, and a variety of non-visual recording methods. An example of the latter would be the use of an event recorder to create a record of the sequence, duration, and frequency of a consumer's activities. An extensive amount of human factors research has been done on human capabilities in the areas of vision, strength, hearing, response time, memory, fatigue, and the effects of environmental factors, such as lighting, temperature, sound, etc. These and other factors are directly related to the consumer's overall ability to accomplish goals without encountering problems.

The evaluation of need solution concepts can be accomplished through the use of the human factors technique of physical simulation. Such a simulation is distinct from engineering test models and product prototypes. A physical simulation may employ none of the technology which will be used in the production product. Instead, it utilizes whatever methods allow the product user to be presented with a realistic approximation of the functional characteristics which the product may exhibit in production form. Ideally, a simulation allows modification of the product characteristics to test a variety of configurations. Simulation provides information on the effects of a product's physical characteristics which are most critical to satisfaction. In addition, the probability of error in the accomplishment of tasks while using the product can be examined. This information can be used both to provide for maximum product safety and provide a basis for defense against liability claims.

Introducing human factors and market research techniques into the new product planning process as discussed, is illustrated in figure 2.

The inevitable question is: "How does the new method work in practice?" In an attempt to provide at least a partial answer, the development of an actual product is outlined below.

A Case In Point: The NurturyTM

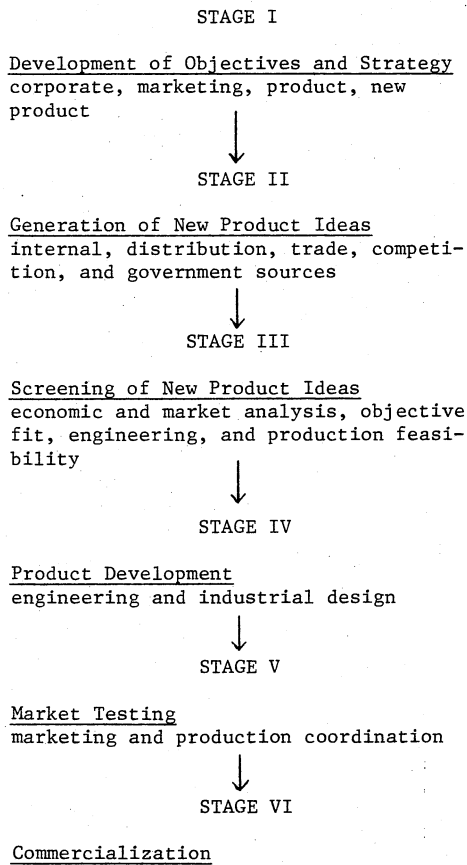
The Nurtury is an infant feeding system which was developed utilizing the process proposed. As a system of products it represents an excellent example of (1) the development of several unique new products, and (2) the development of superior products of a type already on the market. The Nurtury is described in the Appendix.

As discussed earlier in this paper, the first step in the product development process is the identification of consumer need. In the present case, this was accomplished in several ways. First, was the identification of the general area of infant care as a market segment which offered profit potential. An ongoing analysis of basic demographic and social trends in the United States indicated a number of facts which pointed to a potential in this area.

1. A trend to smaller family units which frees disposable income for child care products.
2. Growth of population in the 20 to 39 year age group with an accompanying growth in household formations.
3. An increasing number of first born children in spite of an overall decline in the birth rate.
4. An increase in the average mother's age accompanied by higher educational levels and greater involvement in all facets of child care.
5. The success of firms producing hard goods of the type purchased for first born children in contrast

FIGURE 1

A simplified schematic of the new product planning process. Rejection and feedback loops within and among these stages is implied.



to firms which depended on repeat sales and a high birth rate.

The first step in identifying which of many possible areas within the field of infant care had the greatest development potential was to conduct a series of group discussions with new mothers. Discussions were guided by a moderator from the general topic of infant care problems to specific areas such as feeding, bathing, sleep, mobility, etc. In addition to participating in the discussions, subjects were asked to complete a questionnaire which included listing all infant care products owned and the degree of satisfaction with the products.

The result of these discussions was the identification of home baby food preparation as an area of intense interest for the mothers. The identification of this interest was supported by the finding that 59% of the subjects owned a hand operated baby food grinder which was universally considered an excellent product even though numerous specific problems with its use were cited.

The second step in identifying areas of need in relation to infant feeding was a telephone survey of new mothers to (1) obtain ownership and use information on feeding equipment, (2) determine perceived problem areas, (3) identify preferred characteristics of existing products, and (4) obtain suggestions for improvements in existing products. Each subject was asked specific questions concerning their experiences with the products they

owned. This survey confirmed the need for improved methods of grinding, storing, warming, and serving baby foods. In addition, new areas, such as training aids for drinking and eating were identified.

The final step in the identification of need and the first step in need definition was a literature search. The search served to define existing knowledge concerning the nutritional and safety aspects of home preparation of baby food and provided data on the physical capabilities of infants at various stages of development. The search was supplemented by interviews with pediatricians and pediatric nurses to determine their level of awareness and attitudes toward home preparation. The results of the literature search indicated that underlying the popular interest in home prepared baby food, as revealed by an increasing frequency of publications related to the topic, was a growing body of research which questioned the desirability of feeding commercially prepared foods both from a nutritional and cost standpoint. The chief drawback to home preparation was seen by most professionals as the possibility of contamination during the processing and storage of the food, however, this was not considered an insurmountable problem.

An example of the type of human factors information located concerning physical capabilities of infants was the fact that initial efforts of infants to feed themselves with a spoon are hampered by an inability to perform the combined task of lifting their elbow and rotating their wrist to insert a conventional spoon in their mouth. As a result infants typically invert the spoon while raising it and insert the side of the spoon in their mouth.

Each of the identified areas of need required the application of specific human factors techniques to adequately define exact need characteristics. One of the techniques used to define dimensional requirements for drinking from a training tumbler was an anthropometric survey conducted on a sample population of infants in the defined age group. The survey established the maximum container diameter which could be easily grasped and the most desirable lip design for the tumbler.

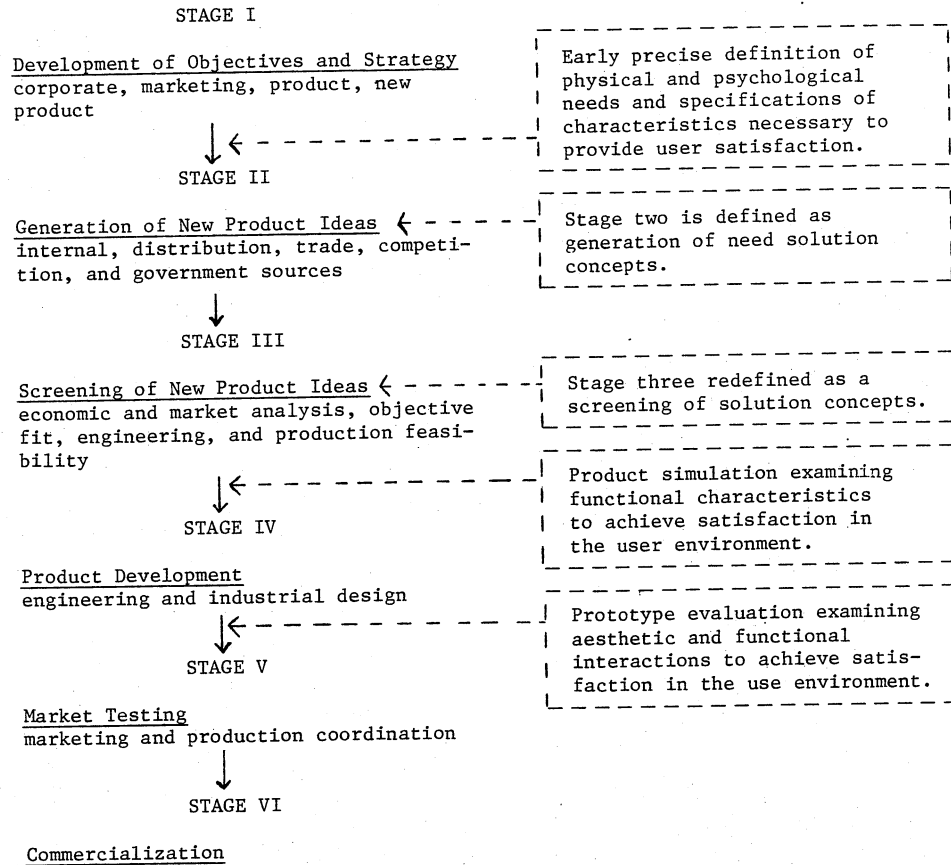
To confirm the information obtained through the literature search and further investigate the existing process of infant feeding a series of observations using slow motion photography were made of mothers preparing and serving baby food in their homes. These films were carefully analyzed to reach an exact definition of problems encountered by individuals using current techniques and equipment.

The problems defined by these studies provided the basis for development of test models or simulations which would allow alternative solution concepts to be realistically evaluated. Each test model was evaluated by fifteen or more mothers and infants for periods of up to two weeks. Subjects were interviewed both before and after the evaluations using detailed questionnaires to define initial perceptions of product characteristics and subsequent changes in perception based on experience with the products. In addition, users were observed and filmed using the test models to aid researchers in the identification of potential problem areas. Depending on the results of the initial evaluations, additional concepts were developed or refinements were made in existing test models. Such modifications were followed by additional user evaluations.

An example of the type of information obtained from observation of the test models in use was the fact that female users of the food grinder invariably applied much greater pressure on the food plunger than was required and as a result often stalled or overloaded the drive motor. This finding indicated that particular attention should be paid to educating the potential user about the

FIGURE 2

INTERJECTING HUMAN FACTORS RESEARCH INTO THE NEW PRODUCT PLANNING PROCESS



proper operation of the product and in addition the grinder should be engineered to meet the actual forces imposed by the user as contrasted with the optimum forces calculated for efficient use.

Conclusion

There are two keys to improving the success ratio of new products. First is utilization of a combination of human factors and market research techniques early in the new product development process to identify and define consumer needs. Second is using these same techniques combined with product simulations to provide consumer input to the screening of solution concepts.

Use of the market research and human factors combine has not been widely exploited and a great deal of refinement and innovation in techniques can be expected. The usefulness of this approach has been demonstrated, however, in the development of a number of products, including those discussed in this paper.

Appendix

The Nurtury™ Infant Feeding System by Teledyne Water Pik is presently being introduced into retail outlets on a national basis. The product consists of three packages containing an electric baby food grinder with accessory items, a baby food warmer and serving tray unit with accessories, and baby feeding helpers including two spoons, a training tumbler, and a bib. Each of the products has characteristics which provide unique

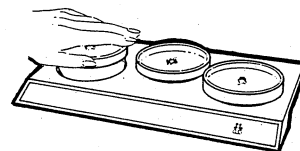
competitive advantages and all of these characteristics were developed to solve problems encountered by mothers and infants in the tasks associated with infant feeding. The primary promotional appeals of the Nurtury are:

1. Mothers can feed their babies the freshest, most natural foods in the world.
2. Babies get the full nutritional value from fresh foods.
3. Mothers save money by turning family food into baby food so there's no waste.
4. Mothers know exactly what's going into their babies.
5. Mothers can easily control the portion size because the Nurtury is specially designed to prepare small quantities of food.
6. Baby food tastes better because it's fresher.

Following are illustrations of each of the products accompanied by promotional copy describing the primary characteristics which make the products unique.

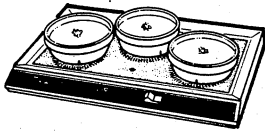
The Nurtury Serving Tray

1. Holds three Storable Servers snugly in place.
2. Has a strong suction cup to firmly grip baby's tray so there's no worry of upsetting.



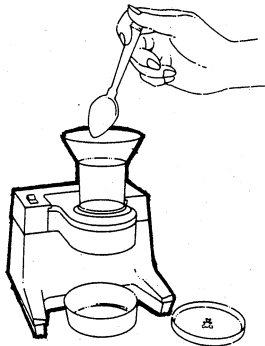
The Nurtury Baby Food Warmer

1. Evenly warms food to the proper temperature in 10-15 minutes.
2. Automatically maintains temperature so food won't overheat or dry out.
3. Used with the same Storable Servers used for preparing food.



The Nurtury Baby Food Grinder

1. Grinds freshly cooked meats, fruits and vegetables into nutritious baby food.
2. Specially designed to prepare meal-size portions so there's no waste.
3. Comes with three texturizing discs for puree, junior, or toddler.



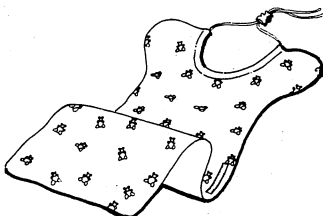
Storable Servers

1. 4-oz., meal-size, resealable containers for preparing, serving, refrigerating and freezing food.
2. Color-coded for easy identification.
3. Transparent, tight sealing lids.
4. Ounce markings on the side to measure meal-size portions directly into Servers.
5. Stackable, easily portable.
6. Made of durable plastic.
7. Freezer-proof to allow preparation and storage of extra food.



The Overall Bib

1. Covers shoulders, chest, sides, and lap.
2. Pocket catches spills, and pulls open for easy cleaning.
3. Made of durable plastic coated nylon.
4. Drawstring neck opens wide to fit over head. Can be adjusted to fit closely and comfortably around neck.
5. Stain-resistant and easy to clean.
6. May be laundered in automatic washer and dryer.



The Nurtury Nutrition Guidebook

A complete guide to infant nutrition including all aspects of introducing solid foods and solving problems encountered.

Mother's Feeding Spoon

1. Bowl has wide, deep portion in front so mother can feed baby without placing entire bowl in baby's mouth.
2. Food stays on tip of spoon for easy feeding. Allows baby to get more food.
3. Bowl is sized for baby's mouth. No sharp or rough edges.



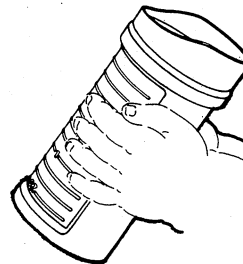
Baby's Self-Feeding Spoon

1. Holes in the bowl help hold the food on the spoon.
2. Holes help to fill the spoon as baby plays in food.
3. Rectangular shaped bowl and special baby-grip handle let baby eat from the sides of the spoon when held in either hand.



Adjusta-Flo Training Tumbler

1. Locking cap won't come off when dropped.
2. Prevents spills because the Tumbler is designed to roll when dropped so holes are at the top when it stops.
3. Metered flow control adjusts to baby's drinking capabilities.
4. Specially designed lip aids baby's natural transition from sucking to drinking, and allows baby to drink without tilting head back.



References

C. Merle Crawford, "Marketing Research and the New Product Failure Rate," Journal of Marketing, 4 (March, 1977), 51-61.

David J. Luck, Hugh G. Wales, and Donald A. Taylor, Marketing Research, (Englewood Cliffs, N.J.: Prentice-Hall Inc., 1970), 447-450.

A. D. Shocker, et. al. Marketing Involvement in Society and the Economy, (Chicago: American Marketing Association, 1969), 167-175.

G. L. Smith and L. B. Archer, "A Methodology for Consumer Design," Proceedings of the Seventeenth Annual Meeting of the Human Factors Society, (Santa Monica, Calif.: Human Factors Society, 1973), 105-110.

Edward M. Tauber, "Why Concept and Product Tests Fail to Predict New Product Results," Journal of Marketing, 39 (October, 1975), 69-71.

INFANT, CHILD AND TEENAGER ANTHROPOMETRY FOR PRODUCT SAFETY DESIGN

Richard G. Snyder, University of Michigan
Lawrence W. Schneider, University of Michigan
Clyde L. Owings, University of Michigan

Abstract

Anthropometry plays an important role in human factors considerations for safe product design, particularly for the 31% of the U.S. population under age 19. This paper discusses two nationwide studies conducted for Consumer Product Safety Commission to obtain center of gravity, linkage, shape and functional body measurement data on 8154 infants, children, and teenagers representative of the U.S. population, for consumer product design, hazard assessment and guidance in establishing requirements or recommendations in standards.

Introduction

For the 77 million individuals under age 19 (Bureau of Census, 1976) representative body dimensional data have been for the most part either non-existent or of little functional value in real-world applications. Yet consideration of body size has a basic relationship to human factors of product safety design. This paper will discuss the data collection problems, techniques, and results of two nationwide anthropometric studies conducted for the Consumer Product Safety Commission (CPSC) by the Highway Safety Research Institute (HSRI), the University of Michigan, from 1972 to 1977 (R.G. Snyder, et al., 1975; 1977).^{1,2,3}

Recent data from the Consumer Product Safety Commission indicate that more than 2 million children are injured each year in accidents involving toys, playground equipment, bicycles, and other child and household products.⁴

The Public Health Service has estimated that toy products alone are responsible for injuries to 700,000 children each year (U.S. Department of Health, Education and Welfare, 1970) and that from 403,000⁵ to 760,000

¹ This study was annually reviewed by, and conducted under standards established by the University of Michigan Medical Center, Committee to review Grants for Clinical Research and Investigation Involving Human Beings, and conforms to the guidelines of the Institutional Guide to Department of Health, Education and Welfare Policy on Protection of Human Subjects.

² This project has been funded with federal funds from the United States Consumer Product Safety Commission under contract number CPSC-C-75-0068. The content of this publication does not necessarily reflect the views of the Commission, nor does mention of trade names, commercial products, or organizations imply endorsement by the Commission.

³ Martha L. Spencer, Herbert M. Reynolds, M. Anthony Schork, and D. Henry Golomb should receive credit for their contributions to one or both of the studies cited.

⁴ Personal Communication, Consumer Product Safety Commission, Bethesda, Md., February, 1975.

⁵ Consumer Product Safety Commission National Electronic Injury Surveillance System (NEISS), 1973.

(Pascarella et al., 1971) children require medical attention each year from bicycle accidents. While a large proportion of these fatalities or injuries are due to causes beyond the control of the product designer, evidence is accumulating to document that many accidents to children are a direct result of poorly or improperly designed commercial products.

A toy or household product may be physically unsafe for many reasons (e.g. sharp edges, points and protrusions, faulty mechanisms, etc.), but a primary factor which has general application to almost all products is concerned with the physical dimensions and features of the product and its components as they relate to the physical characteristics of the user or other individuals who may come into contact with the product. For example, how large or small should holes and spacings in toys, furniture slats, railings, and appliance ventilation covers be so that infants and children do not get their hands, fingers, and heads stuck or injured? How should strollers, high chairs, and tricycles be designed to prevent accidents due to instability? Or, what should be the dimensions of bicycles and bicycle hand brakes, school desks, school bus seats, playground equipment, etc. designed for use by children of specific ages and/or body sizes? Answers to these and many other questions require that a complete and reliable source of the physical characteristics of infants, children, and youth be available to designers, engineers, and other groups concerned with establishing guidelines and standards for the manufacture of these products.

An example of the type of anthropometric data required to answer functional design problems is demonstrated by the question of crib slat spacing. For three to six-month old infants what distance between slats will safely restrain them? Since no applicable data were available in the literature, in 1972 an infant population sub-study was designed to take 11 measurements (Snyder, et al. 1972). It was determined that the most critical body measurement was that of buttocks depth, measured from the maximum protrusion of the buttocks to the anterior surface of the upper legs. Findings from the study, which also included limited *in-situ* tests, indicated that the expected 90% of the population at a confidence level of 95%, lies between 2.9 inches and 4.7 inches for 4½ to 6½ month old infants and 2.9 inches to 4.4 inches for 2½ to 4½ month old infants. These were viewed as conservative as they were taken with pressure transducers on soft tissue in a relatively non-compressed condition. Since inter-slat distances, particularly on older cribs, were found to be as high as 4 inches, it was obvious that a design change was necessary. The CPSC subsequently established the standard as 2-3/8 inches (Federal Register, 1973).

As of 1972 a source of reliable and consistent data providing measurement statistics for both traditional and functional anthropometric measures for a sample representative of the infants, children, and youth in the United States was not available. This was surprising in view of the fact that some 800 studies in the literature were found to provide child measurement data (Snyder et al., 1973; reprinted 1975). This comprehensive review of the literature on child and infant anthropometric data revealed that while a large number of studies had been

conducted, the type of measurements needed for most safety applications were actually non-existent.

A large number of these studies were longitudinal (rather than cross-sectional) surveys designed to study the growth of a fixed population over a period of time using only a few basic measures such as height and weight. All the studies were limited in that they either used specific populations, selected age ranges, or only a few measurements, and did not consider differences due to region of the country, race, and other environmental and socio-economic factors likely to affect the sizes of children. Even if these factors had been considered and if all the measurements necessary had been taken in one study or another, differences due to measurer technique, definitions in measurements, and differences in age groupings would make it impossible to compile these data into one common source of reliable and consistent information. Additionally, many growth studies as recently as Marshall and Carter (1975) have found that children today are taller and larger than children their same age in past generations. Since two-thirds of the 35 most complete studies encountered had been made prior to 1960, it is questionable how representative much of these data may be of the current population of U.S. children.

This point was emphasized in a 1975 study conducted jointly by physical anthropologists from several organizations attempting to evaluate masterbody forms for 3-year and 6-year old child dummies for the Federal Aviation Administration (Young et al, 1975) and National Highway Traffic Safety Administration (Reynolds et al, 1975). It was found that of 98 body dimensions determined to be necessary for describing the external morphology of a child test manikin, 68 were not available in the literature. For example, the shape of the torso at the nipples is described by circumference, breadth, depth, and height from seat pan, but only circumference and depth measurements had been previously taken. Unfortunately this work was conducted before data were available from the initial CPSC Study. Problems related to anatomical considerations for the design of child restraints, another application where anthropometric data were needed, has been outlined by Burdi et al (1969).

Objectives

The purpose of the two studies conducted for CPSC was to provide anthropometric data of both a functional and traditional nature to describe the sizes, shapes, body linkage, selected functional measures, and seated and standing centers of gravity for a population of infants, children, and teenagers representative of the 48 conterminous United States. Tasks involved determining the measurements most required; the experimental design; designing, constructing, and testing unique advanced anthropometric instrumentation including mini-computer systems and entirely new center of gravity devices; collection, reduction, and analysis of the required dimensional data; and finally, tabulation and representation of the data in a format most useful to persons concerned with designing or establishing guidelines and standards for children's products, taking into consideration the potential functional applications.

Methods

This study was undertaken as a multi-disciplinary team effort with co-investigators representing specialties in physical anthropology, bioengineering, pediatrics, computer engineering, industrial and operations engineering and biostatistics. In addition, advisory consultation with other scientists occurred as the study progressed.

To establish the numbers of various types of infants and children to be included in the sample, the initial study

utilized U.S. Census and HEW data to project a sample of 4000 subjects whose racial and socio-economic characteristics matched the proportions of those characteristics in the total U.S. population. While the 1972-1975 study involved 41 measurements on infants and children from age 2 weeks to 13 years, the 1975-1977 study expanded this data base to also include 13 to 18 year olds (a population for which little data were known), and expanded the number of measurements to 87.

To assure the national representativeness of school-age subjects, a sample of schools was selected from a national list available from the third round of a study of youth fitness completed in the Department of Physical Education at the University of Michigan. The selection process used in the youth fitness study has been described in detail in Reiff, et al. (1968). The U-M Survey Research Center maintains demographic and socio-economic data on a national sample of geographic-political units representative of the U.S. population. Utilizing this sampling frame as a base, a smaller sampling frame was conducted for the anthropometric study. It consisted of random selection of eight central-city areas, 10 suburban areas, and 31 areas outside large metropolitan areas. The next step involved random selection of one school district within each of the sampled areas. Finally, specific schools in those districts were randomly selected, using a ratio of two elementary schools to one intermediate school to one senior high school. Because of scheduling and space problems, some schools originally selected did not participate in the measurement study. However, every effort was made to select matching schools from the same area.

Measurements and Measurement Strategy

Traditionally, anthropometric measurements have been devised to describe human variation in body morphology, body proportions, and changes in size attributed to growth, race, and other variables of interest to anthropologists. These dimensions were taken to describe the linear distance between two landmarks (heights, breadths, and depths) or around a body segment at a prescribed level (circumference). In today's world, however, this classical approach to anthropometry is not sufficient to answer questions relevant to the interaction of man with a multitude of man-made environments. While the traditional anthropometry measures are still important as they describe differences between individuals and populations, the differences must also be considered as a part of a man-machine system. Thus a systems approach to anthropometry was used as a general guiding principle. Consideration was also given to the probability that another study of this magnitude would not be likely in the near future and it was thus essential to provide the maximum amount of information possible within the budget and time limitations and to identify those measurements that would be of most interest and usefulness to the greatest number of potential users.

The result of these considerations was a list of 87 measurements shown on the following page. Traditionally, all measurements are made on every subject in the survey sample. However, because time available per subject was limited to 15-20 minutes so as to assure school and subject cooperation, this permitted obtaining only about 40 measurements per subject. The strategy adopted was to obtain the 22 Core measurements (Group I) on all subjects, plus Group II, Group III, or Group IV measurements sequentially, on every third subject. Thus, 42-45 measurements were taken on each subject, and the sample sizes for measurements in Groups II, III, and IV are one-third as large as for Group I.

The key factor in this approach was to select measurements for Group I that correlate highly with measurements in the other three groups. Because the core measurements were taken on every subject, and each of the

non-core measurements correlates highly with one or more of the core measurements, this provided a means of checking the representativeness and comparability of the non-core sample populations. Statistical comparisons of measurement means in selected core measurements and age groups showed that there is no significant difference between the samples for Groups II, III, and IV.

To avoid any systematic association of measurements with subjects in a particular age group or region of the U.S. all measurement sets were taken equally at all measurement sites. To accomplish this, the combination of measurement sets (I + II, I + III, I + IV) were done sequentially.

Measurements, Age 2-19

I. Core Measurements. Weight, Stature, Head Circumference, Chest Circumference, Waist Circumference, Hip Circumference, Upper Arm Circumference, Forearm Circumference, Upper Thigh Circumference, Calf Circumference, Foot length, Foot Breadth, Hand Length, Head Breadth, Shoulder Breadth, Shoulder-Elbow Length, Elbow-Hand Length, Maximum Hip Breadth (sitting), Buttock-Knee Length, Knee Height, Erect Sitting Height, Head Breadth.

II. Shape Measurements. Chest Height, Waist Height, Hip Height, Chest Breadth, Waist Breadth, Neck Circumference, Natural Waist Circumference, Wrist Circumference, Ankle Circumference, Calf Circumference, Ankle Breadth, Calf Breadth, Upper Thigh Depth, Wrist Breadth, Forearm Depth, Upper Arm Depth, Neck Breadth, Head Length, Eye Height, Thigh Clearance, Maximum Thigh Breadth (sitting).

III. Linkage Measurements. Vertical Arm Reach, Frontal Arm Reach, Lateral Arm Reach, Step Height, Suprasternale Height, Iliocristale Height, Iliospinale Height, Trochanteric Height, Gluteal Furrow Height, Sphyrion Height, Tibiale Height, Bispinous Breadth, Clavicale-Acromion Length, Acromion-Radiale Length, Radiale-Styilion Length, Biacromial Breadth, Hip Breadth at Trochanter, Supine Stature, Supine Sitting Height, Standing Center of Gravity, Seated Center of Gravity.

IV. Head, Face, and Hand Measurements. Bizygomatic Breadth, Maximum Frontal Breadth, Mouth Breadth, Nose Length, Lower Face Height, Face Height, Head Height, Tragon to Back of Head, Tragon to Top of Head, Ear-Sellion Depth, Bitragon Breadth, Ball of Fist Circumference, Maximum Fist Breadth, Maximum Fist Depth, Thumb-Crotch-Middle Finger Length, Middle Finger-Thumb Grip Length, Thumb Length, Index Finger Length, Middle Finger Length, Middle Finger Diameter, Index Finger Diameter, Thumb Diameter, Minimum Hand Clearance.

Infant Measurements. Since infants (age range two weeks to 24 months) do not have mature skeletal and muscular systems and cannot perform movements required for most of the functional measurements, a separate set of 34 measurements was compiled and used for these subjects. These were:

Weight, Crown-Sole Length, Crown-Rump Length, Head Circumference, Head Breadth, Head Length, Shoulder Breadth, Shoulder-Elbow Length, Upper Arm Circumference, Elbow-Hand Length, Forearm Circumference, Wrist Circumference, Hand Length, Hand Breadth, Minimum Hand Clearance, Maximum Fist Breadth, Thumb Diameter, Middle Finger Diameter, Chest Circumference, Chest Breadth, Waist Circumference, Waist Breadth, Rump-Sole Length, Rump-Knee Length, Hip Circumference, Hip Breadth, Mid-Thigh Circumference, Mid-Thigh Depth, Knee-Sole Length, Calf Circumference, Ankle Circumference, Ankle Breadth, Foot Length, Foot Breadth.

Data Collection

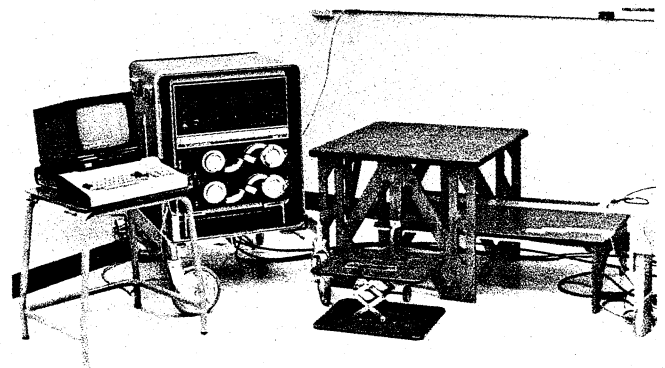
In the 1972-1975 study, measurements were made at 71 schools, clinics, and day care centers in 8 states. The 1975-1977 study included measurements obtained at 104 locations in eight regions (20 states) by two teams. In the latter study, one consisted of two males, the other of two females. Team members were selected for their ability and experience in working with children and teenagers. They received extensive training and supervised practice before commencing the field work. Several members were fluent in Spanish which was important in some schools particularly in the Southwest where English was a second language. In some cases materials to parents were printed in both Spanish and English. The process of obtaining approval from school authorities and parents, and of scheduling specific grades and classes, usually required considerable preliminary contacts. Ordinarily the school superintendent was approached by letter and then telephone. The request was presented to the school board for approval and then principals of schools in that district were contacted. Often the team coordinator visited the schools to explain the study to teachers and students, using an 18-minute film (University of Michigan, 1975) produced during the previous CPSC-sponsored study. After it was decided that a school would participate, an explanatory letter, consent form, and demographic questionnaire were distributed to parents of children in school grades selected to participate. The entire process of obtaining approval, informing teachers, parents, and students, and obtaining signed consent forms usually required about two months. When schools were closed during summer vacations, subjects were measured at summer camps.

The primary sources for subjects two to five years old were nursery schools and day care centers. Nationwide, infants were the most difficult age group to obtain, and additional sources for infant subjects were The University of Michigan Well Baby Clinic and a pediatric clinic in Pontiac, Michigan.

Instrumentation

An important aspect of this study was the design, fabrication and use of a portable mini-computer system for automatic retrieval of measurement data. Figure 1 illustrates this system and shows the NOVA 1220 computer system including 24 K of memory, 2 LINC type drive units 16 channels of A/D input, a signal conditioning and power supply package for processing instrument signals and activating device transducers, and an interactive display system consisting of keyboard, terminal and video display.

Figure 1



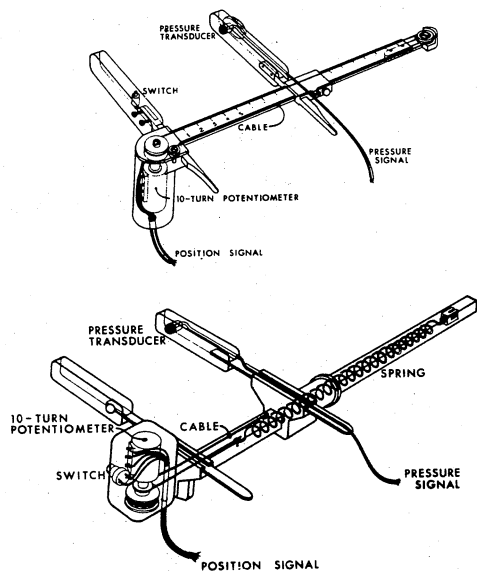
The Automated Anthropometric System

Four different automated measuring devices were interfaced with the NOVA computer system and utilized in the measuring process. These have been reported in detail by Snyder, et al. (1972; 1975), and Owings et al. (1974 a; 1974 b; 1974 c).

While computerized techniques have been suggested and explored previously (Garn, 1962; Prah1-Anderson, 1972; Bullock, 1974⁶), the automated anthropometric system described is believed to be the most extensive systematic use of automated anthropometry, the first such use of the NOVA mini-computer data acquisition system, and the first practical portable computerized means of obtaining center of gravity measures on both infants and children.

For measuring linear distances (i.e., breadths, lengths, depths, heights) the automated caliper or anthropometer was used. These instruments are illustrated in the drawings of Figure 2. GPM (Siber Hegner & Co. Ltd) anthropometers and calipers were modified to provide electrical readout of length by means of a 10-turn potentiometer connected to the moving blade by a pulley and

Figure 2



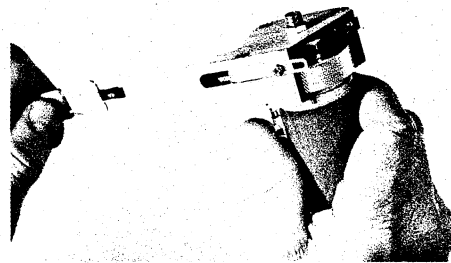
Modified sliding calipers and modified anthropometers showing potentiometer, pulley and cable system, and pressure transducer.

cable system. A miniature pressure transducer (Königsberg Model P21) mounted in the special plexiglass blade provided a means of standardizing measurements on soft tissue. Body linkage and head, face, and hand measurements were taken using removable pointed extensions of different lengths. A button in the stationary plexiglass blade activates the measuring process on each device.

A third automated device shown in Figure 3 was used for measuring various body circumferences. This consists of a standard steel measuring tape wrapped around a pulley which is attached to the shaft of a 10-turn potentiometer. The tape is wrapped about the particular body segment being measured and clipped back on the tip of the aluminum tape guide. Tension is provided by a coil spring located inside the pulley and the measurement is sensed by the computer upon depressing the plastic but-

ton with the thumb or index finger.

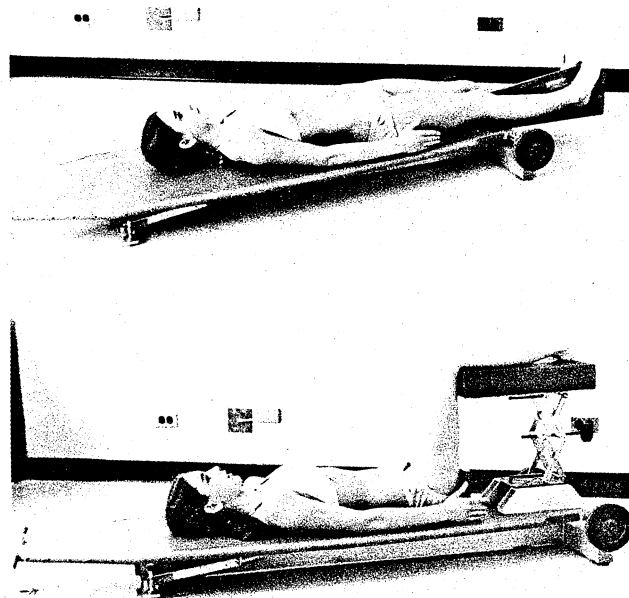
Figure 3



Modified girth device for electrical readout of circumference measurements.

Two devices were designed for measuring whole body center of gravity, one for infants and one for children and teenagers up to 250 pounds. The principal of operation is the same for both and is illustrated for the larger device in Figure 4. The subject is placed in either the supine standing or sitting position on a rigid platform supported by three precisely calibrated load cells. Outputs of the load transducing cells are amplified before entering the A/D unit of the computer. By sensing the load cell readings with the subject in position and with the device empty, the computer calculates the location of the center of gravity relative to a known reference plane.

Figure 4



Center of gravity device with subject in supine standing (upper) and supine sitting position (lower).

Two complete systems including computer, measurement instruments, TV display, and center of gravity devices were assembled and transported throughout the United States by the two teams in Dodge Maxi-vans as illustrated in Figure 5.

Other devices used for taking other functional measurements include a hole template for measuring finger and thumb diameters, hinged plexi-glass hole boards for measuring hand clearance (Figure 6) and a grip rod with pivoting extension for measuring functional grip dis-

⁶ Margaret I. Bullock, University of Queensland, Brisbane, Australia (personal communication, 1974)

tances (not shown). In order to standardize and facilitate the positioning of the subjects for various measurements, a portable table and adjustable foot rest

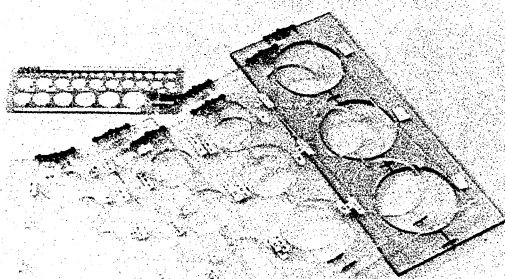
Figure 5



Maxi-van used to transport anthropometric equipment.

platform were designed and used with each measuring system (See Figure 1).

Figure 6



Hole template for measuring finger diameter and hinged plexiglass hole boards for measuring hand clearance.

The entire measuring process including sequencing through the measurement data sets, calibration of automated instruments and pressure transducers, recording of subject demographic information, retrieving and recording measurement values from the various automated devices, checking measurement values relative to subject age and stature, and calculating subject center of gravity was controlled by a computer program called "MAP" or Michigan Anthropometric Processor.

Data Handling, Editing, and Analysis

At the completion of each geographic location data tapes were returned to HSRI for transfer to a file on the Michigan Terminal System (MTS). Interim editing took place during the course of the study by listing measurement values according to subject age and visually checking for bad data. Final editing was accomplished when data collection was complete by using two routines written for the HSRI PDP 11/45 system.

A program called X-VAL, obtained from anthropologists at Wright Patterson AFB, was used to search for extreme measurement values in designated age groups. Subjects

were grouped by data set, sex, and age intervals of one year each, and the program listed the ten smallest and ten largest values for each measurement along with statistics of the distribution (e.g., mean, standard deviation, skewness). Visual inspection of the data was then used to search for subject numbers with unreasonable measurement values and a complete listing of the data on these subjects was produced. From this listing a decision could be made as to whether to delete or leave the value in question.

A second procedure used to edit the data was multiple regression analysis. This involved establishing a linear relationship between one measurement variable and several other measurement variables. For example, one might attempt to establish a linear relationship of weight with stature, chest circumference, and waist circumference [weight = f (stature, chest circ., waist circ.)]. Once a relationship was established, measurement values which fell outside the predicted value, plus or minus specified error tolerances, were edited appropriately. This procedure allowed for editing points which were incorrect but which still fell within two standard deviations of the mean measurement value for a given age range (i.e., they are not extreme values).

Sample Population Characteristics

The composition of the sample population of measurement subjects by race and sex is shown in Table 1. As the table shows, 50.7 percent of the subjects were male, 49.3 percent were female. As for ethnic identification, 86.5 percent of the subjects were identified as White, 11.0 percent as Black, .9 percent as Oriental, .1 percent as American Indian, and 1.5 percent as Other (mixed or unknown).

Every effort was made to ensure the ethnic, geographic, and socioeconomic representativeness of the sample. However, biases resulting from self-selection factors due to the voluntary nature of final subject measurements were recognized.

Table 1

Race	Value	Male	Female	Total
White	n	1817	1753	3570
	row %	50.9	49.1	100.0
	col %	86.9	86.1	86.5
Black	n	225	227	452
	row %	49.8	50.2	100.0
	col %	10.8	11.2	11.0
Oriental	n	18	19	37
	row %	48.6	51.4	100.0
	col %	0.9	0.9	0.9
American Indian	n	5	1	1
	row %	83.3	16.7	100.0
	col %	0.2	0.0	0.1
Other	n	27	35	62
	row %	43.5	56.5	100.0
	col %	1.3	1.7	1.5
Total	n	2092	2035	4127
	row %	50.7	49.3	100.0
	col %	100.0	100.0	100.0

Sample Population by Race and Sex

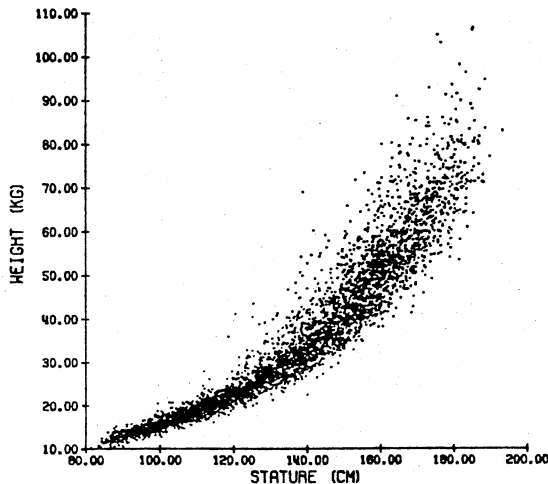
Results

The anthropometry and center of gravity measurements for 8,154 infants, children, and teenagers were obtained, collated, and analyzed in these two nationwide CPSC studies. For detailed information concerning the resulting data the basic reports (Snyder et al, 1975; 1977) must be referred to. The final report of the most recent study presents the measurement results for 34 measurements of infants from two weeks to 24 months old, and for 87 measurements of children and teenagers from two through 18 years old. For each measurement, tables present the summary statistics for males, females, and combined sexes in 16 age groups, including sample size (N), mean, standard deviation, minimum measurement, 5th 50th, and 95th percentiles, and maximum measurement. Also presented in tables and graphs are the bivariate relationships for selected functional measurements with stature and weight. The linear regression coefficients and constants given in these tables provide for predicting these functional measurements on individuals or groups for which only stature and weight are known.

To illustrate the results and functional nature of the measurements, results for five of the measurements are presented in the Appendix. These measurements in order of presentation are stature, weight, frontal grip reach, distance from thumb crotch to tip of the middle finger, and seated center of gravity as a percent of sitting height.

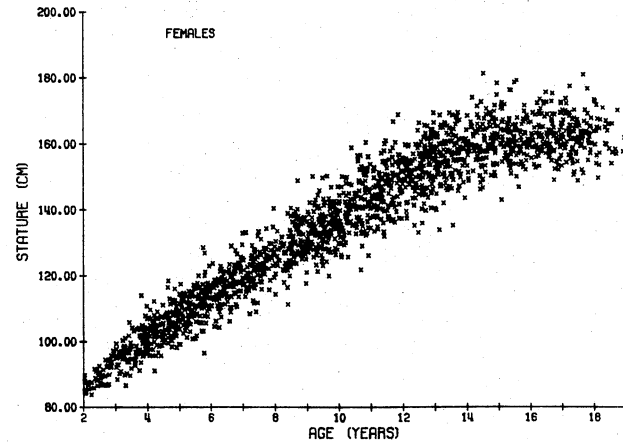
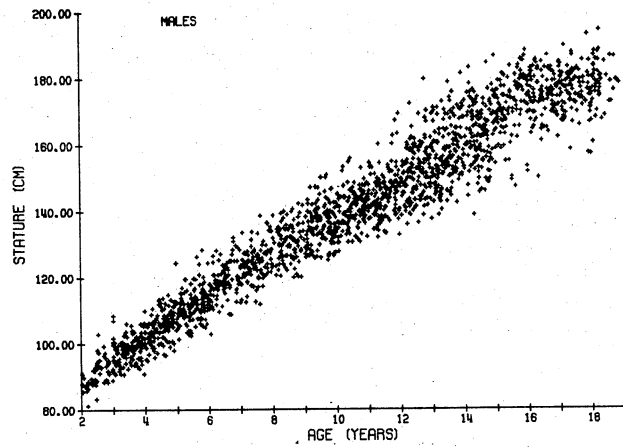
It is worth noting that since publication of the final reports several new and diverse applications of these data have been made known to us. For example, requests for these data have been made by a foreign government relative to child restraints, a major aerospace company relative to wrist dimensions, a manufacturer of medical x-ray phantoms for facial data, a U.S. assistance program for application to nutritional problems, several toy manufacturers, an automotive manufacturer relative to child dummies, a group concerned with design of school bus seats, and several other government agencies. While no single study can fulfill all the needs of all users it is the authors' hope that this body of new data made possible by the major technical advances in instrumentation of the automated computerized anthropometry system, will provide the user with basic functional, shape, and linkage data necessary for improved product safety design.

Appendix



STATURE

Subject stands erect with head oriented in the Frankfort Plane, arms hanging at sides. With an automated anthropometer, measure the vertical distance from the standing surface to vertex (top of the head).



STATURE (cm)
(Males)

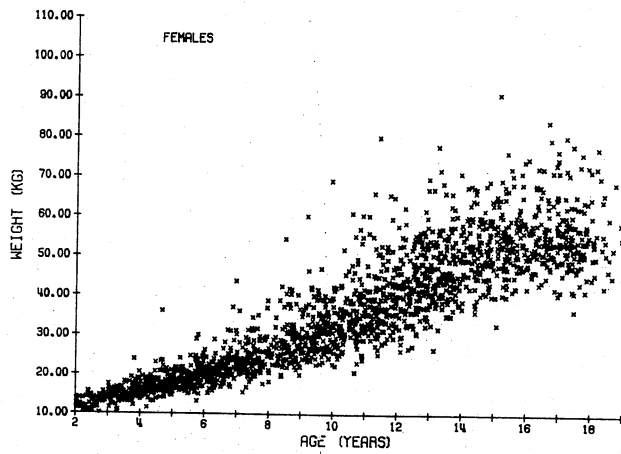
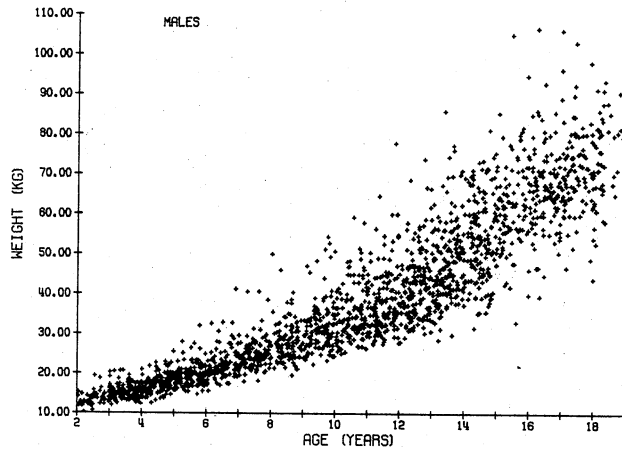
Age (yrs)	N	Mean	s.d.	Min	5th	50th	95th	Max
2.0-3.5	115	94.5	5.0	81.3	87.0	94.3	102.2	108.5
3.5-4.5	118	101.3	4.5	90.9	93.8	100.8	109.0	113.3
4.5-5.5	144	108.6	4.7	96.5	100.6	108.9	115.5	124.4
5.5-6.5	117	115.1	5.2	102.9	106.1	114.6	124.7	128.4
6.5-7.5	165	122.0	5.1	108.9	112.6	122.5	130.6	133.8
7.5-8.5	104	127.8	5.6	111.9	119.9	126.7	137.7	140.0
8.5-9.5	117	133.4	6.1	120.0	123.8	133.4	142.4	147.3
9.5-10.5	124	137.9	6.3	125.2	127.0	137.6	149.2	155.9
10.5-11.5	142	142.5	5.3	130.7	133.8	142.1	151.1	160.2
11.5-12.5	154	148.4	7.4	132.8	138.1	147.4	161.3	172.4
12.5-13.5	154	154.1	8.4	136.4	142.1	153.4	169.5	179.8
13.5-14.5	155	161.3	8.7	138.8	147.3	160.6	175.5	193.2
14.5-15.5	131	166.4	8.7	146.0	149.7	167.0	179.6	186.7
15.5-16.5	100	174.5	7.8	147.3	161.1	174.6	185.2	189.3
16.5-17.5	104	175.9	6.0	158.4	165.3	176.2	184.3	189.6
17.5-19.0	88	177.1	6.8	157.1	166.2	176.9	187.4	194.4

STATURE (cm)
(Females)

Age (yrs)	N	Mean	s.d.	Min	5th	50th	95th	Max
2.0-3.5	98	92.1	4.7	83.8	85.1	92.0	99.4	105.9
3.5-4.5	110	101.6	4.6	91.1	93.9	101.7	108.7	114.1
4.5-5.5	127	108.0	4.6	98.2	100.5	108.0	116.0	118.6
5.5-6.5	125	114.2	5.1	96.5	104.9	114.3	121.7	128.7
6.5-7.5	122	120.5	5.7	106.3	110.5	120.4	129.5	133.6
7.5-8.5	94	125.9	5.5	111.5	116.4	125.8	134.5	140.6
8.5-9.5	140	132.7	5.9	117.8	124.1	132.3	142.2	150.3
9.5-10.5	134	137.5	6.2	120.1	127.0	137.3	147.3	159.0
10.5-11.5	140	144.2	7.6	122.0	133.0	144.0	156.3	161.1
11.5-12.5	133	149.3	7.0	133.3	137.3	149.3	159.9	163.0
12.5-13.5	161	155.1	7.0	134.1	142.0	155.6	165.7	169.6
13.5-14.5	116	158.4	7.0	135.4	146.8	158.2	169.0	174.3
14.5-15.5	133	162.0	6.4	143.2	152.3	161.3	172.4	181.6
15.5-16.5	98	162.1	6.1	147.1	153.2	161.2	172.6	179.5
16.5-17.5	117	162.5	5.9	149.1	152.3	162.6	171.6	177.7
17.5-19.0	68	163.0	5.9	149.9	153.0	163.0	172.7	181.3

WEIGHT

Subject stands on clinical scales dressed in briefs or bathing suit. Weight is measured to the nearest tenth of a kilogram and typed into the computer.



WEIGHT (kg)
(Males)

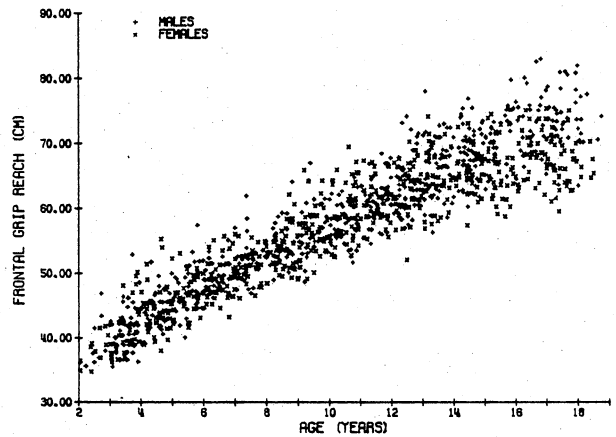
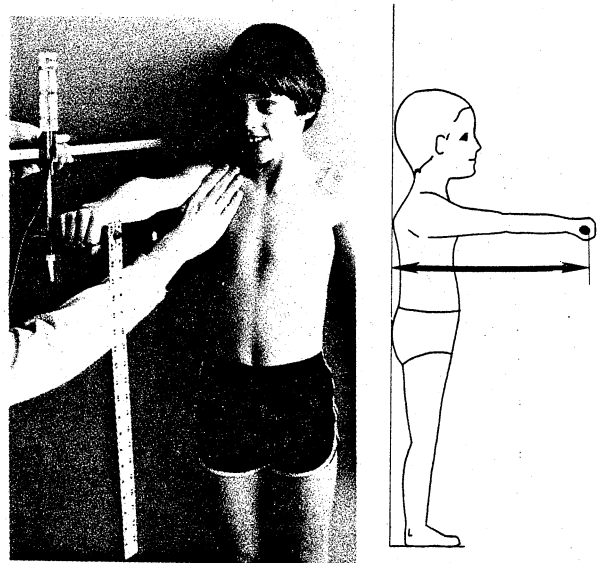
Age (yrs)	N	Mean	s.d.	Min	5th	50th	95th	Max
2.0-3.5	115	14.6	2.0	10.2	11.9	14.4	18.9	20.7
3.5-4.5	118	16.2	2.1	12.3	13.1	15.8	19.7	22.7
4.5-5.5	144	18.4	2.3	13.7	15.2	18.2	22.6	27.9
5.5-6.5	117	20.8	3.1	14.8	16.8	20.1	26.6	32.5
6.5-7.5	106	24.0	3.9	17.4	18.9	23.0	29.6	41.2
7.5-8.5	104	26.9	5.0	18.2	20.9	25.7	35.3	43.9
8.5-9.5	116	29.8	5.1	19.6	22.7	29.0	37.9	48.3
9.5-10.5	124	33.5	6.7	23.3	25.1	32.0	47.9	54.3
10.5-11.5	142	36.6	7.2	26.3	27.8	34.4	50.3	54.5
11.5-12.5	154	40.0	7.9	27.8	30.3	38.1	54.7	77.8
12.5-13.5	154	45.4	9.6	28.5	33.1	43.7	63.6	85.8
13.5-14.5	155	51.8	9.6	29.8	38.0	50.9	69.6	77.3
14.5-15.5	131	57.5	11.2	35.6	39.5	56.6	74.9	105.0
15.5-16.5	100	66.1	11.9	33.0	46.2	66.1	85.7	106.6
16.5-17.5	104	69.2	10.5	46.8	54.2	67.4	88.5	106.2
17.5-19.0	88	73.2	11.0	43.9	55.2	72.2	91.1	112.3

WEIGHT (kg)
(Females)

Age (yrs)	N	Mean	s.d.	Min	5th	50th	95th	Max
2.0-3.5	97	13.6	1.7	10.1	10.8	13.7	16.3	17.5
3.5-4.5	110	16.2	2.2	11.4	13.0	16.0	19.4	23.8
4.5-5.5	127	18.2	2.6	14.0	14.9	17.7	21.6	36.0
5.5-6.5	126	20.2	2.9	13.1	15.7	19.9	24.6	29.9
6.5-7.5	125	23.4	4.3	15.4	17.5	22.8	31.2	43.6
7.5-8.5	93	26.2	5.5	17.6	19.4	25.1	37.3	54.2
8.5-9.5	140	29.6	5.9	20.0	22.5	28.2	40.4	50.0
9.5-10.5	134	32.6	6.5	22.2	24.2	31.9	42.2	63.0
10.5-11.5	140	37.7	9.1	20.7	26.5	35.5	55.3	30.0
11.5-12.5	133	40.8	8.2	26.1	28.3	40.0	54.6	55.5
12.5-13.5	161	46.1	8.5	26.5	34.1	44.8	59.7	77.8
13.5-14.5	116	49.5	8.5	34.9	36.8	49.0	66.7	71.0
14.5-15.5	133	53.5	8.2	32.7	40.5	52.7	67.2	91.0
15.5-16.5	98	55.1	7.9	41.4	43.0	53.3	70.7	74.8
16.5-17.5	117	55.9	8.6	36.4	41.9	54.2	73.6	84.1
17.5-19.0	68	55.9	8.4	41.4	43.8	53.9	72.7	77.3

FRONTAL GRIP REACH

Subject stands erect with feet together, back to wall, grasping the handle of the grip device in right hand. The anthropometrist holds the subject's right shoulder against the wall as the subject extends right arm to maximum horizontal grip reach. With the pointed blade of an automated anthropometer, measure the horizontal distance from the wall to the most distal point on the handle of the grip device.

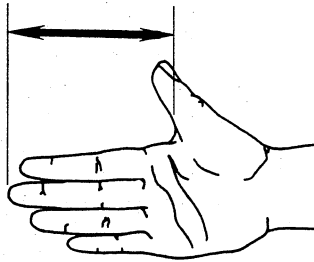
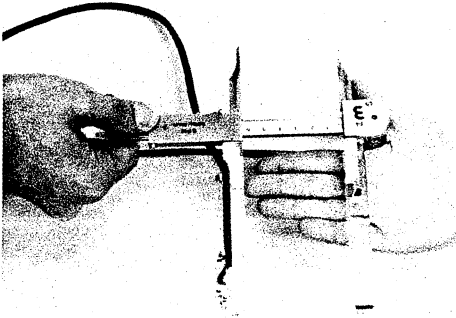


FRONTAL GRIP REACH (cm)
(Males and Females)

Age (yrs)	N	Mean	s.d.	Min	5th	50th	95th	Max
2.0-3.5	62	39.7	3.1	34.7	35.5	39.0	45.4	48.1
3.5-4.5	77	43.1	3.6	36.3	37.8	42.5	49.7	52.9
4.5-5.5	74	45.4	3.4	38.0	40.4	45.2	51.5	55.3
5.5-6.5	76	48.5	2.9	41.6	44.2	48.1	53.3	57.4
6.5-7.5	74	50.7	3.2	43.2	46.2	50.7	56.1	61.9
7.5-8.5	64	52.9	2.7	46.5	47.8	53.1	56.8	59.3
8.5-9.5	80	55.5	4.1	48.6	49.2	55.1	64.1	67.0
9.5-10.5	74	57.7	3.3	52.5	52.8	56.3	64.0	66.0
10.5-11.5	94	59.9	3.8	51.7	53.4	59.9	66.6	69.5
11.5-12.5	91	62.1	3.8	52.1	56.3	61.5	67.7	74.2
12.5-13.5	97	64.9	4.5	56.3	57.6	64.3	71.6	78.1
13.5-14.5	80	67.0	4.3	57.4	59.4	66.6	74.1	77.0
14.5-15.5	87	67.7	3.8	59.2	61.3	67.6	73.4	75.6
15.5-16.5	63	69.6	5.1	58.8	60.9	63.9	76.4	80.2
16.5-17.5	74	70.6	5.0	59.6	62.0	70.7	78.7	83.1
17.5-19.0	44	70.7	5.4	62.8	63.0	70.2	80.3	92.1

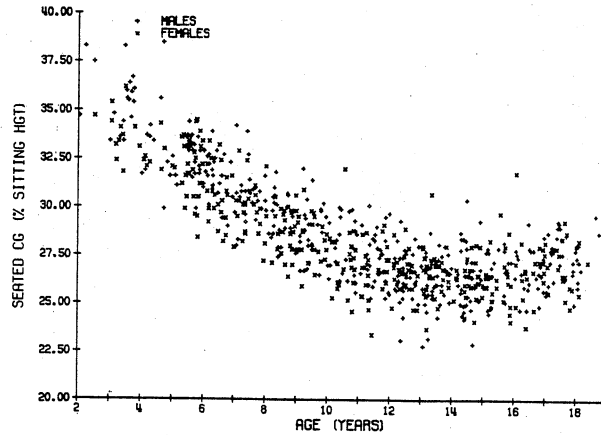
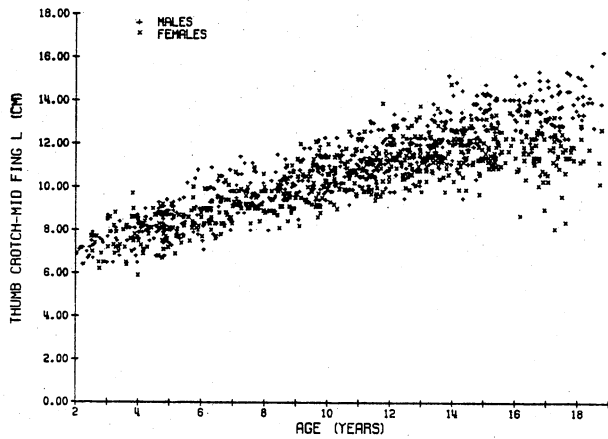
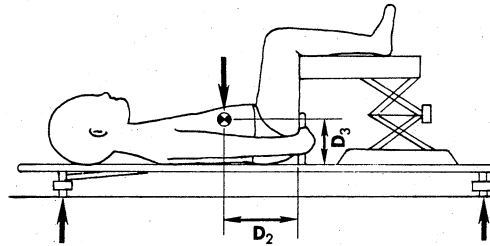
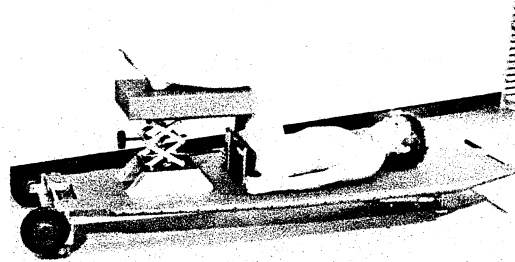
THUMB CROTCH-MIDDLE FINGER LENGTH

Subject extends right hand and fingers with palm up, thumb away (abducted) from hand. With an automated sliding caliper, measure the distance from the crotch of the thumb to the tip of middle finger parallel to the fingers. A constant correction factor for the caliper blade offset is added to the measurement.



SEATED CENTER OF GRAVITY

Subject lies on back on the CG device with legs placed over the adjustable support so that the knees form a 90° angle, the buttocks are firmly against the reference plane, and the upper legs are approximately 90° to the torso. The arms are placed at the sides. The horizontal distance of the center of gravity from the buttocks reference plane (D_2) and vertical distance of the center of gravity from the CG platform (D_3) are computed. The horizontal distance is reported as measured and as a percent of supine sitting height [$(D_2/\text{supine sitting ht}) \times 100$]. The vertical distance is also reported as measured and as a percent of buttock-knee length [$(D_3/\text{buttock-knee L}) \times 100$].



THUMB CROTCH-MIDDLE FINGER LENGTH (cm)
(Males and Females)

Age (yrs)	N	Mean	s.d.	Min	5th	50th	95th	Max
2.0-3.5	54	7.4	0.6	6.2	6.4	7.2	8.5	8.8
3.5-4.5	62	7.9	0.7	5.9	6.5	8.0	8.8	9.7
4.5-5.5	83	8.2	0.7	6.7	6.8	8.1	9.2	9.5
5.5-6.5	59	8.7	0.8	7.1	7.5	8.5	10.2	10.9
6.5-7.5	62	9.4	0.8	7.8	8.2	9.3	10.7	11.4
7.5-8.5	60	9.5	0.8	8.0	8.1	9.4	10.8	11.5
8.5-9.5	79	10.0	0.8	8.2	8.8	9.9	11.3	12.0
9.5-10.5	92	10.6	0.9	8.6	9.1	10.5	12.0	12.3
10.5-11.5	98	11.0	0.8	9.2	9.7	10.8	12.5	13.1
11.5-12.5	94	11.3	1.0	9.4	9.7	11.2	12.9	13.9
12.5-13.5	107	11.7	0.8	9.8	10.2	11.6	13.1	13.9
13.5-14.5	97	12.0	1.1	9.5	10.2	12.0	13.8	15.2
14.5-15.5	95	12.5	1.1	9.7	10.7	12.3	14.2	14.8
15.5-16.5	55	12.7	1.3	8.7	10.5	12.6	14.4	15.2
16.5-17.5	80	12.4	1.4	8.1	10.1	12.4	14.7	15.4
17.5-19.0	46	13.0	1.6	8.4	10.3	12.9	15.4	16.3

SEATED CG (% SITTING HEIGHT)
(Males and Females)

Age (yrs)	N	Mean	s.d.	Min	5th	50th	95th	Max
2.0-3.5	19	34.6	1.8	31.8	**	34.2	**	38.3
3.5-4.5	21	34.2	1.6	31.7	31.7	33.9	36.4	36.7
4.5-5.5	27	32.6	1.8	29.9	29.9	32.2	35.1	38.5
5.5-6.5	65	31.9	1.4	28.4	29.3	32.1	33.9	34.5
6.5-7.5	48	30.5	1.6	27.9	28.0	30.4	33.1	34.2
7.5-8.5	44	29.5	1.3	27.1	27.2	29.7	31.5	32.1
8.5-9.5	56	28.7	1.3	25.9	26.5	28.7	30.7	32.0
9.5-10.5	41	27.9	1.4	25.3	25.8	27.8	29.9	30.2
10.5-11.5	52	27.2	1.6	23.4	24.7	27.1	29.8	32.0
11.5-12.5	64	26.9	1.3	23.1	24.9	26.7	29.1	29.7
12.5-13.5	64	26.4	1.3	22.8	23.9	26.4	28.0	30.7
13.5-14.5	57	26.5	1.3	24.2	24.3	26.4	28.4	30.4
14.5-15.5	59	26.2	1.1	22.9	24.5	26.1	28.0	29.3
15.5-16.5	42	26.6	1.6	23.8	24.1	26.6	29.1	31.8
16.5-17.5	44	27.2	1.1	24.7	25.0	27.2	29.0	29.3
17.5-19.0	32	27.1	1.3	24.8	24.9	26.9	28.9	29.6

References

Alphonse R. Burdi, Donald F. Huelke, Richard G. Snyder, and George H. Lowrey, "Infants and Children in the Adult World of Automotive Safety Design: Pediatric and Anatomical Considerations for Design of Child Restraints," Journal of Biomechanics, 2(1969), 267-280.

"Requirements for Full-Sized Baby Cribs," Title 16, Chapter 2, Subchapter C, Part 1508, Federal Register, 38 (November 21, 1973), 32129-32132.

Stanley M. Garn; "Automation in Anthropometry," American Journal of Physical Anthropology, 20(1962), 387-388.

W. A. Marshall, and C. O. Carter, "Child on Mid-Parent Regression for Full Adult Stature," Annals of Human Biology, 2 (Fall 1975), 307.

Clyde L. Owings, Lawrence W. Schneider, Richard G. Snyder, and Martha L. Spencer, "A Portable System for Infant and Child Center of Gravity Measurement" Proceedings, Conference on Engineering, Medicine and Biology, (October, 1974), 375.

"Computer Controlled Anthropometry: A Portable System for Use with Infants and Children" Proceedings, Conference on Engineering, Medicine and Biology, (October, 1974), 385.

Clyde L. Owings, Richard G. Snyder, Martha L. Spencer, and Lawrence W. Schneider, "New Techniques for Infant and Child Anthropometry: Mini-Computer Controlled Anthropometry and Center of Gravity Measurements," American Journal of Physical Anthropology, 41(November, 1974) 497.

E. A. Pascarella, J. P. Foley, D. Levine, and J. R. Steward, Characteristics of Youthful Bicycle Riders in an Urbane Community and Events Occurring to Operation of Their Vehicles (Chapel Hill: University of North Carolina Highway Safety Research Center, June, 1971).

B. A. Prah-Andersen, J. Pollman, D.J. Roaben, and K. A. Peters, "Automated Anthropometry," American Journal of Physical Anthropology, 37(1972), 151-154.

Guy G. Reiff, Leslie Kish, and Jean Harter, "Selecting a Probability Sample of School Children in the Coterminous United States," American Association for Health, Physical Education, and Recreation, Research Quarterly, 39 (May, 1968), 409-414.

Herbert M. Reynolds, Joseph W. Young, John T. McConville, and Richard G. Snyder, Development and Evaluation of Masterbody Forms for Three-Year Old and Six-Year Old Child Dummies (U.S. Department of Transportation, National Highway Traffic Safety Administration, Washington, D.C. DOT HS-801811, January, 1976.

Richard G. Snyder, Martha L. Spencer, Clyde L. Owings, Lawrence W. Schneider, and M. Anthony Schork, Selected Infant Anthropometry Crib Slat Sub-Study (Prepared for Consumer Product Safety Commission, Bethesda, Md. Final Report UM-HSRI-BI-74-4, Highway Safety Research Institute, The University of Michigan, Ann Arbor, Michigan.

Richard G. Snyder, Martha L. Spencer, Clyde L. Owings, and Peter Van Eck, Source Data of Infant and Child Measurements: Interim Data. (Prepared for Children's Hazards Division Bureau of Product Safety, Food and Drug Administration, Bethesda, Md. The University of Michigan, Ann Arbor. January, 1973; reprinted January, 1975).

Richard G. Snyder, Martha L. Spencer, Clyde L. Owings, and Lawrence W. Schneider, Physical Characteristics of Children as Related to Death and Injury for Consumer

Product Safety Design. (Prepared for Consumer Product Safety Commission, Bethesda, Md. Final Report UM-HSRI-BI-75-5, Highway Safety Research Institute, The University of Michigan, Ann Arbor, Michigan. 1975 NTIS PB-242221 \$7.50). Also published as, Anthropometry of U.S. Infants and Children (Society of Automotive Engineers, Inc., Warrendale, Pa. Special Publication SP-394, Paper No. 750423, February, 1975).

Richard G. Snyder, Lawrence W. Schneider, Clyde L. Owings, Herbert M. Reynolds, D. Henry Golomb, and M. Anthony Schork, Anthropometry of Infants, Children, and Youth to Age 18 for Product Safety Design (Prepared for Consumer Product Safety Commission, Bethesda, Md. Final Report UM-HSRI-77-17, Highway Safety Research Institute, The University of Michigan, Ann Arbor, Michigan. May, 1977. NTIS PB-270227 \$16.25).

University of Michigan, Highway Safety Research Institute "New Methods of Infant and Child Anthropometry for Product Safety" (Produced for Consumer Product Safety Commission, Bethesda, Md. March 21, 1975. Film, 18 minutes, color).

U.S. Bureau of the Census, Statistical Abstract of the United States: 1976. (97th edition) Washington, D.C., 1976, 27.

U.S. Department of Health, Education, and Welfare, Injury Control Program, Estimates of Injuries from Consumer Products. Submitted at National Commission on Product Safety Hearing of October 21, 1969.

Joseph W. Young, John T. McConville, Herbert M. Reynolds, and Richard G. Snyder, Anthropometric Dimensions Representative of Average Three and Six-Year Old Children Sizes for the Construction of Mastermodel Body Forms (U.S. Department of Transportation, National Highway Traffic Safety Administration, and Federal Aviation Administration, Washington, D.C., Oklahoma City, Okla. DOT HS-801638. April, 1975).

SAFETY RESEARCH: YOU AND YOUR LAWN MOWER

A.M. RameySmith, National Bureau of Standards
V.J. Pezoldt, National Bureau of Standards
J.J. Persensky, National Bureau of Standards

Abstract

The Human Factors Section at the National Bureau of Standards was recently involved in the evaluation of a Lawn Mower Safety Standard for the Consumer Product Safety Commission. This report describes the results of three studies performed in conjunction with this effort as related to blade contact injuries and the implementation of a dead-man control requirement.

Introduction

The Human Factors Section at the National Bureau of Standards (NBS) has provided support to the Consumer Product Safety Commission (CPSC) during the development and evaluation of a power lawn mower safety standard. Part of this support took the form of providing CPSC with data concerning how lawn mower users interact with power mowers. Three empirical studies were performed to assist CPSC in resolving specific questions concerning requirements in the proposed safety regulation based on a report from Consumers Union (Consumers Union, 1975).

In October 1974, CPSC accepted the offer of Consumers Union of United States, Inc. (CU) to develop a standard under Section 7 of the Consumer Product Safety Act [15 U.S.C. 2056(b)]. The CPSC received a proposed standard from Consumers Union in July 1975 (CU, 1975). In May 1977, CPSC published in the Federal Register (CPSC, 1977) a proposed safety standard based on the CU proposal with modifications and additions made by the CPSC staff.

CPSC estimated that there were 178 000 power lawn mower-related injuries in 1975. Costs associated with these injuries are estimated to be in excess of 73 million dollars. According to the National Electronic Injury Surveillance System, operated by CPSC, over half of the injuries resulted from body contact with the lawn mower blade (CPSC, 1976). In an attempt to reduce the number of blade contact injuries, the CPSC included two requirements in the proposed safety regulation which directly relate to blade contact.

The first requirement, intended to protect feet from blade contact, describes a region under the lawn mower housing through which the blade cannot pass. This is accomplished by using an anthropometric foot probe in a performance test. The second requirement is for a "dead-man" control which must be continuously activated by the mower user in order for the blade to operate. The "dead-man" control is primarily intended to protect hands and fingers from blade contact. The three studies reported here, all performed during the evaluation phase of the standard development process, deal with these two requirements.

The first experiment addresses the problem of describing the region under a mower housing through which the blade should not pass. This was accomplished by a comparison of three available foot simulator probes with a sample of potential lawn mower operators' feet, especially with regard to dynamic characteristics.

The second experiment provided data to answer the question "How long after the dead-man control is released should the blade be allowed to rotate?" Alternatively, one might ask "How long does it take an operator to move from the handle location to an area of potential contact with a moving blade?"

The third experiment is related to the prospect that consumers might intentionally defeat the dead-man control. To meet the proposed standard the blade may be stopped either by means of a blade clutch/brake system with the engine continuing to run or by engine shut-down. If this latter method is employed to stop the blade, lawn mower users would likely be required to restart their mowers repeatedly throughout a period of use. The CPSC has suggested that this inconvenience may encourage consumers to defeat the dead-man control unless the mower is easy to restart. This experiment, therefore, was designed to provide CPSC with data which can be used to objectively define the subjective judgment of easy-to-restart for pull-started lawn mowers.

Complete technical details and data from the three studies can be found in the NBS reports on the individual studies (Ramey and Persensky, 1977; Pezoldt and Persensky, 1977; Persensky and Ramey, 1977).

Study I. Evaluation of Anthropometric Foot Probes

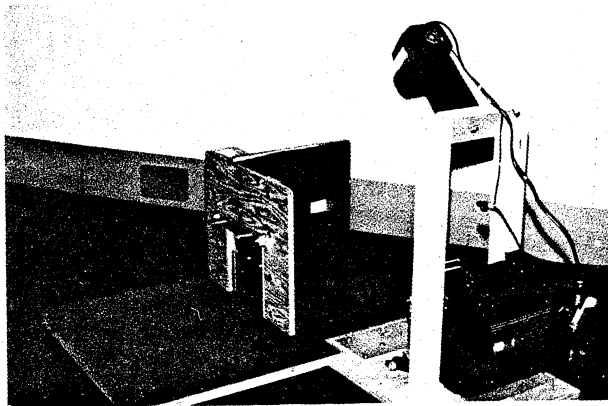
The foot probe test developed by CU for inclusion in the proposed mandatory lawn mower standard makes two major assumptions: first, that the probe or foot simulator is representative of the feet of the population of lawn mower users; and second, that the probe test determines potential for foot contact with the blade.

NBS evaluated the adequacy of available generic foot probes, emphasizing a comparison between the probes and the dynamic characteristics of the feet of potential lawn mower users. The probes used in the study were: the Consumers Union (CU) proposed probe, adapted from an Underwriters' Laboratory probe; the United Kingdom (UK) probe, designed by the British Standards Institute; and the American National Standards Institute (ANSI) probe.

The apparatus used in the study (Figure 1) consisted of two parts. First a camera and strobe were mounted on a platform focused on a gray-on-black measured-graph backdrop. The film used produced positives for immediate examination and negatives for future analysis and was capable of multiple exposure. The second part of the apparatus was a frame with a movable metal partition 15 cm wide which could be adjusted to allow vertical openings of 6 to 10 cm in 1 cm increments to simulate different lawn mower housing heights.

The subjects in the study included 127 males and 74 females ranging in age from 9 to 66 years. Foot length data for these subjects were compared to those collected by other researchers for specific populations. This comparison demonstrated that the present sample was representative of the range of potential operators' foot sizes. No determination however, could be made of the representativeness of the distribution of foot sizes.

FIGURE 1
EVALUATION OF ANTHROPOMETRIC
FOOT PROBES TEST APPARATUS



After putting a white nylon stocking (to increase image contrast) over their right foot, subjects were instructed to:

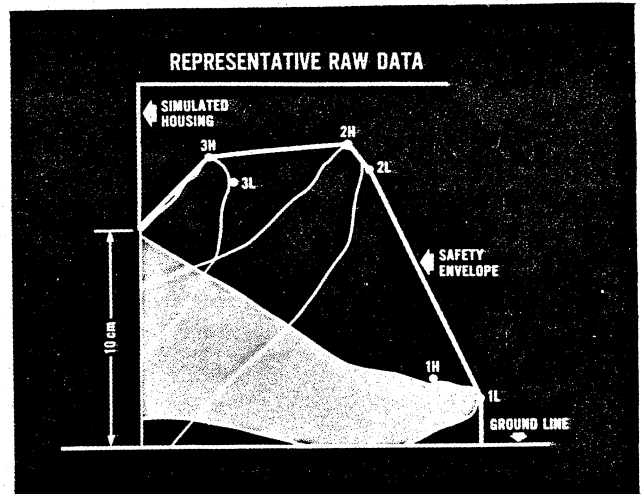
Insert their right foot under the metal panel as far in and to the left as comfortable, keeping the foot flat on the base. Then,

Rock back on the heel of the foot, lifting the front of the foot as far off the base as possible, curling toes upward, yet keeping the foot as far in as possible.

Keeping the front of the foot elevated and toes curled up, pull the foot out from under the partition until the base of the toes contacted the metal panel.

At each of the three positions described above a picture was taken, creating a triple exposure print and negative as graphically depicted in Figure 2.

FIGURE 2
EVALUATION OF ANTHROPOMETRIC FOOT
PROBES REPRESENTATIVE RAW DATA

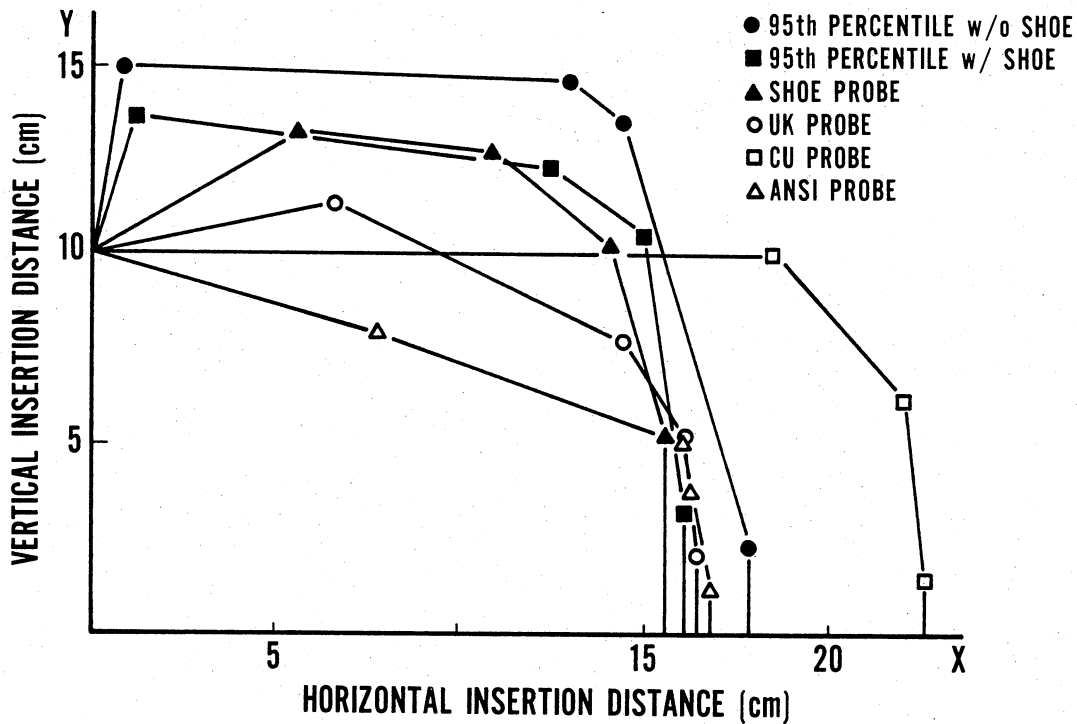


Triple exposure photographs, one at each of the three housing heights (i.e., 6, 8 and 10 cm), were taken both with shoes on and shoes off. This procedure resulted in six triple exposure photographs, one exposure at each of the three positions described above, for each participant.

The negative photographs were used for data reduction. Measures were taken for the point representing the height of the longest point of each of the three positions and the point corresponding to the length of the highest point at each position. (Position 1 was foot flat on floor; Position 2 was foot raised and toes curled; and Position 3 was base of toes contacting the partition.) There were, therefore, two sets of coordinates for each position, or a total of six per negative (housing height). The same procedure was repeated for each participant at each housing height both with shoe on and shoe off--a total of 36 sets of coordinates per participant. The same three positions were measured for each of the three generic foot probes and for a shoe probe made by affixing a shoe on the end of a length of wood. The shoe probe was constructed from a 95th percentile size female's shoe.

Horizontal and vertical insertion distances for the generic probes, the shoe probe, and the foot data were used to develop safety envelopes for each housing height. Figure 2 shows the measurement points used to develop the safety envelopes. Figure 3 is a composite envelope at a housing height of 10 cm for each of the three generic probes (CU, UK, ANSI), the shoe probe, and the 95th percentile foot data.

FIGURE 3
 COMPOSITE SAFETY ENVELOPE FOR THREE GENERIC
 PROBES, A SHOE PROBE AND THE 95TH
 PERCENTILE FOOT DATA AT 10 cm HOUSING HEIGHT



Of particular interest is the fact that, of the three generic probes, only the UK probe passes through the space above the plane of the housing opening. However, none of the probes approaches the vertical insertion distance achieved by the 95th percentile foot data. This region, above the housing, is of primary importance in providing foot protection since mower blades are located above the plane of the bottom of the housing. Although the CU probe exceeds the other generic probes and the foot data in terms of horizontal insertion distance, it does not provide any additional foot protection over that provided by the UK probe since the mower blade rotates above the lower edge of the housing.

In addition to comparing safety envelopes derived from the aggregate foot data with the generic probe data, a comparison of each individual's data was made with each generic probe to determine the percentage of participants who would be completely protected by each generic probe and by the shoe probe. These data indicate that at least one point of each individual's foot movement data falls outside the safe area defined by each of the generic probe envelopes. Therefore, a lawn mower which satisfies the criteria of any of the generic foot probes would not completely protect any of the participants in study. The shoe probe performed somewhat better than the generic probes. However, as shown in Table I, this probe would not

provide the magnitude of protection which would be obtained from a probe designed from the 95th percentile foot data.

TABLE I
 PERCENTAGE OF SAMPLE COMPLETELY PROTECTED
 BY SHOE PROBE AND BY A PROBE DESIGNED
 TO 95TH PERCENTILE FOOT DATA

Sample Condition	Housing Height (cm)		
	6	8	10
% Protected by Shoe Probe			
With Shoe	14	14	46
Without Shoe	1	1	14
% Protected by 95th Percentile Probe			
With Shoe	77	82	74
Without Shoe	80	78	73

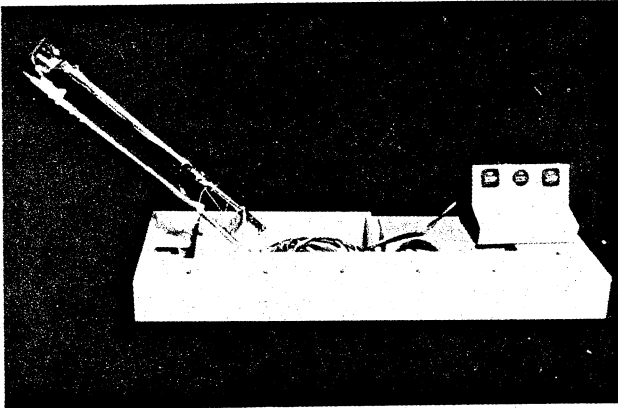
Clearly, there is a serious problem with all of the existing generic probes, as shown by comparison of the individual data with the generic probe data. It seems clear that further research should be performed toward the development of a more adequate probe.

Study II. Time-to-Blade-Access

Several recommendations have been made for the maximum allowable blade stopping time, that is, the time from release of the dead-man control until the blade comes to a complete stop. The power mower industry, through the Outdoor Power Equipment Institute (OPEI), suggested a stopping time of seven seconds (OPEI, 1975). Consumers Union recommended a three or four second stopping time based on a study employing college students (CPSC, 1977). The National Bureau of Standards had previously recommended a blade stopping time of one second, based on data reported by the University of Iowa (Porter, Jones, and Persensky, 1974). In view of the inconsistency among these recommendations, CPSC requested NBS to perform a basic empirical study of operator movement time to provide an ergonomically sound recommendation for blade stopping time for walk-behind power mowers.

Of critical interest is the time it takes a lawn mower user to move from the operating position at the handle to a position of potential blade contact. Participants in this study were tested using a reaction time device which was designed to measure mower users' approach times to various points representing areas of potential blade contact. The test apparatus (Figure 4) permitted measurement of reaction time--or time to release a simulated dead-man control at the onset of a cue light and, more importantly, movement time--or the time from the release of the dead-man control to activation of one of five switches as surrogates for blade access areas.

FIGURE 4
TIME-TO-BLADE ACCESS TEST APPARATUS



The subjects' task was to move from the handle to the appropriate switch designated by a cue light, then depress the switch by hand. The distances of the switches from the handle were at intervals of 25.0 cm from 50.0 cm to 150.0 cm. The shortest and longest distances were based on the distance of the closest rear wheel and furthest front wheel, respectively, as measured on a sample of existing lawn mowers. The 100 subjects tested (64 males and 36 females) ranged in age from 16 to 62 years. Each received five trials at each distance on a randomized schedule. Subjects were instructed not to rush to strike the switch plates, but rather to walk at a normal rate.

Blade-area access-time data are summarized in Table II. Analyses of these data reveal that, as expected, there are no statistically significant differences in reaction time as a function of movement distance; however, movement time increases with movement distance. Average movement times observed in this study ranged from 0.6 sec. to 3.3 sec. The median movement times at the shortest and longest distances were 1.4 and 2.2 sec., respectively. Clearly, however, a "worst case" strategy seems more fitting when recommending an appropriate blade stopping time. That is, the selected time should protect significantly more than half of all users. Table II also presents percentile information for movement time to each of the five simulated blade contact areas. Inspection of the table suggests that a blade-stopping time of 0.7 second would protect all but five percent of the sample from potential contact with the rotating blade at the shortest distance tested.

TABLE II
TIME-TO-BLADE-ACCESS: REACTION
AND MOVEMENT TIMES (SEC)

	Movement Distance (cm)				
	50.0	75.0	100.0	125.0	150.0
Reaction Time					
Median	0.5	0.5	0.6	0.6	0.6
Range	0.3-1.2	0.3-2.0	0.3-1.1	0.3-1.3	0.3-2.0
Movement Time					
Median	1.4	1.5	1.8	2.0	2.2
Range	0.6-2.7	0.6-2.8	0.7-3.0	0.8-3.3	0.9-3.0
5th Percentile	0.7	0.8	0.9	1.1	1.2
95th Percentile	2.2	2.5	2.8	3.0	3.2

The time data generated in this study reflect only direct movement time from the handle to areas of potential blade contact. In the absence of any data to the contrary, however, it must be assumed that in the worst case lawn mower operators do, in fact, move directly from the operating position to an area of potential blade contact. Unless other movement scenarios can be established, a more valid movement time criterion cannot be determined.

Study III. Ease of Pull

This study was designed to provide CPSC with data based on subjective judgments, which can be used to objectively define "easy-to-restart" for pull-started power lawn mowers. Many factors are involved in establishing criteria, among the more important of which are:

- the force of pull required to start the engine;
- the distance through which a pull must be made;
- the average number of pulls required per engine restart;
- the number of times the engine must be restarted during the mowing period; and
- the time interval between restarts.

In addition to these factors, which are all external to the individual starting the mower, a number of human characteristics are of at least equal importance. To the obvious factors of age, sex and physical condition, must be added the kinesthetic and proprioceptive feed-

back cues experienced when pulling the starting cord and the connotations placed upon the terms "easy" and "hard."

This study did not attempt to explore all of these factors. The data generated by this effort, therefore, cannot be construed as providing a definitive answer to the question "What is easy to restart?" Rather, the present study attempts to define easy to pull under experimentally controlled conditions. In psychological terms, the problem becomes one of determining the relationship between physical stimuli and the psychological responses to such stimuli. In this case the stimuli are the forces which are exerted on a simulated pull-start mechanism and the responses are subjective judgments about the ease or difficulty involved in applying these forces.

Two simulators were designed and built for use in this study (Figures 5 and 6). The simulators were operationally identical with the exception of the locus of pull. Pulls on the "housing" simulator were executed from a height typical of several current lawn mower designs, i.e., 38 cm (15 in). The "handle" simulator was designed to be pulled from a position on the handle, 83 cm (33 in) above the ground. Although most current lawn mowers are more similar to the housing simulator, lawn mowers incorporating a dead-man control may have the pull handle located in an area similar to that on the handle simulator.

FIGURE 5
EASE OF PULL TEST APPARATUS
HOUSING SIMULATOR

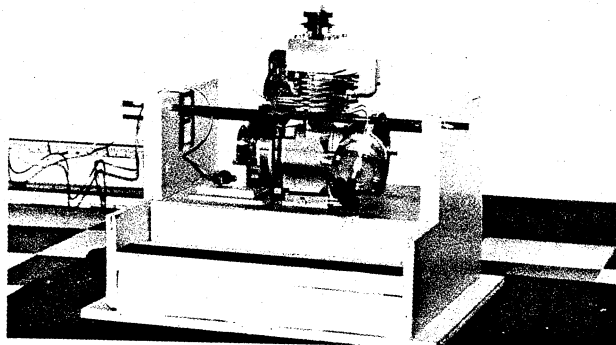


FIGURE 6
EASE OF PULL TEST APPARATUS
HANDLE SIMULATOR



The simulators permitted pulls similar to those experienced when pull-starting a power lawn mower. Pulls could be made through four distances and the difficulty (resistance) of pulls could be manipulated by the experimenter. Difficulty of pull could only be controlled in a very gross sense by manipulating an adjustable cylinder on the simulator engine; the actual force applied was determined by the participant. Measurement of the forces applied when pulling the start cord was accomplished by incorporating a load cell in the pull handles. These load cells were interfaced with a digital readout, peak load indicator. Peak force (i.e., the maximum force exerted at any point in the pull) was measured to the nearest pound (4.45N). Forces required to actually start the engine could not be determined since neither gasoline nor spark plugs were present in the simulators.

The subjects' task consisted of making 36 pulls at one of the simulators on each of four consecutive days. Each day the subjects were provided with a different length cord which determined the distance through which they were to pull. The distances through which they pulled were approximately 46, 61, 76, and 91 cm (18, 24, 30, and 36 in). The order in which subjects were assigned to the various distance categories was randomly determined. Subjects were instructed to pull the "starting" cord in the same manner they would to start a lawn mower. Immediately following each pull, subjects were requested to judge whether that pull was easy or hard. A forced choice paradigm was employed, requiring subjects to judge each pull as either "easy" or "hard." No intermediate responses were allowed. Participants were instructed to consider the simulators as "perfect lawn mowers," that is, they would start on every pull. In fact, of course, the engine never started. Subjects made six pulls in a five minute test period, followed by a rest period of approximately 10 minutes. Six test periods were completed each day by each subject.

A modified staircase method was employed to determine compression ratio adjustments to the simulator after each pull. Generally, if a subject judged a pull "easy", the compression ratio of the simulator engine was increased one unit for the next pull. If a pull was judged "hard," the compression ratio was decreased. Some variation in this procedure occurred in an attempt to assure that subjects pulled over a broad range of forces. After each pull, the judgment about the pull, either "easy" or "hard," and the peak force exerted during the pull were recorded. Peak force, rather than any measure of force over time, was employed for several reasons. First, peak force proved simple to measure. Second, pilot tests showed the measure to be capable of discriminating between pulls judged easy and pulls judged hard. Finally, peak force was felt to be the most easily adaptable to test method development.

A total of 5,372 pulls were made from the handle position and 5,009 pulls from the housing. For most analyses, the peak force measures of primary interest were the maximum forces judged easy (Easy_{max}) and the minimum forces judged hard (Hard_{min}) for each subject. For purposes of analysis, peak forces were grouped in 22N (5 lbf) categories.

With one exception, there were no statistically significant differences for either Hard_{min} or Easy_{max} as a function of the distance of pull. The peak forces, therefore were combined for all four pull distances. Tables III and IV show the cumulative distributions (number and percent) of Hard_{min} and Easy_{max} for males and females using the handle and housing simulators respectively. These data are shown graphically in Figure 7 for female subjects using the handle simulator.

Figure 7 provides the basis for defining "easy to pull" under the conditions of this study. The cumulative percentages of Hard_{min} and Easy_{max} displayed in this figure define the lower limit of "hard to pull" and the upper limit of "easy to pull" respectively for females pulling from the handle position. Similar figures, not presented here, have been constructed for males using the handle simulator and for males and females using the housing simulator. For any given percentage of the sample, the Hard_{min} line defines the minimum peak force which was judged hard. Similarly, the Easy_{max} line defines the maximum peak force which was judged easy for any specified proportion of the sample. These two functions may be viewed as providing conservative (Hard_{min}) and liberal (Easy_{max}) definitions of the peak forces considered to be easy to pull. Inspection of Figure 7 reveals the very large difference between the "liberal" and "conservative" estimates of easy to pull. This difference reflects the large variability in judgments of easy and hard both between individuals and within a single individual. This variability was evident for males and females using both the handle and housing simulators.

It was recommended that a "conservative" stand with regard to determining a force value for easy to pull be adopted by CPSC. This was suggested so that the greatest percentage of the population could "easily" restart lawn mowers and so that the incentive for defeating the dead-man control would be reduced. In practice, this would mean determining the peak

force by using the Hard_{min} distribution for females and selecting the sample percentage at 80 percent or greater. That is, the value for pulls from the handle position should be approximately 140N (31 lbf).

This study does not, nor was it intended to, answer the question of why one judgment is "easy" and another "hard." Neither does it provide a definitive answer to the question of what is easy to restart. The data generated in the study do, however, provide a practical basis upon which a policy decision regarding the upper force limit for "easy to restart" can be based.

TABLE III
CUMULATIVE DISTRIBUTIONS OF HARD_{MIN} AND EASY_{MAX}
HANDLE SIMULATOR, ALL PULL DISTANCES COMBINED

Female Subjects					
Hard _{min}			Easy _{max}		
Peak Force N	lbf	Cumulative Number	Cumulative Percent	Cumulative Number	Cumulative Percent
98	22	20	100		
120	27	19	95		
142	32	15	75		
165	37	8	40	20	100
187	42	3	15	18	90
209	47			16	80
231	52			14	70
254	57			11	55
276	62			8	40
298	67			5	25
320	72				
343	77			4	20
365	82				
387	87			1	5

Male Subjects					
Peak Force N	lbf	Cumulative Number	Cumulative Percent	Cumulative Number	Cumulative Percent
120	27	18	100		
142	32	15	83		
165	37	12	67		
187	42	11	61		
209	47	8	44		
231	52	7	39	18	100
254	57	4	22	14	78
276	62	3	17		
298	67				
320	72			13	72
343	77	2	11	10	56
365	82				
387	87				
409	92			9	50
432	97	1	6	6	33
454	102			5	28
476	107				
498	112				
521	117			3	17
343	122				
565	127			2	11
587	132			1	6

TABLE IV
CUMULATIVE DISTRIBUTIONS OF HARD_{MIN} AND EASY_{MAX}
HOUSING SIMULATOR, ALL PULL DISTANCES COMBINED

Summary

Female Subjects

Peak Force N	lbf	Hard _{min}		Easy _{max}	
		Cumulative Number	Cumulative Percent	Cumulative Number	Cumulative Percent
120	27	18	100		
142	32	17	94		
165	37	16	89		
187	42	15	83		
209	47	12	67		
231	52	9	50		
254	57	5	28		
276	62	4	22	18	100
298	67			17	94
320	72			15	83
343	77	2	11	14	78
365	82	1	6	11	61
387	87			9	50
409	92				
432	97				
454	102			7	39
476	107			4	22
498	112			2	11
521	117			1	6

Male Subjects

Peak Force N	lbf	Cumulative Number	Cumulative Percent	Cumulative Number	Cumulative Percent
142	32				
165	37				
187	42	16	89		
209	47				
231	52	15	83		
254	57	10	56		
276	62	9	50		
298	67	7	39	18	100
320	72	4	22		
343	77			16	89
365	82	2	11	15	83
387	87			14	79
409	92	1	6	13	72
432	97				
454	102			12	67
476	107			11	61
498	112			8	44
521	117			7	39
543	122			6	33
565	127			4	22
587	132			2	11
610	137				
632	142				
654	147			1	6

The three studies reported here provide empirically based answers to specific questions which have arisen during the development of a proposed safety standard for power lawn mowers. As a result of these studies, the following recommendations have been made to CPSC:

1. In lieu of development of a new foot probe, the United Kingdom probe should be used for testing compliance with the standard since, of available probes, the UK probe most closely approximates movement of human feet under a lawn mower housing.
2. Based solely on the elapsed time from release of the dead-man control to finger access to areas of potential blade contact, a maximum of 0.7 seconds should be allowed for blade stopping.
3. The peak force necessary to restart a lawn mower with a handle mounted pull cord should not exceed approximately 140N if restarting is to be considered easy by a substantial portion of lawn mower users.

All of these recommendations have been made with certain qualifications. However, the three studies reported here do provide data which should be considered when CPSC makes policy decisions in these matters.

References

Consumers Union of United States, Inc. Proposed Safety Standard for Power Lawn Mowers, Submitted to the U.S. Consumer Product Safety Commission, July 1975.

Outdoor Power Equipment Institute, Technical Comments and Recommendations on Proposed Lawn Mower Safety Standard, Submitted to the U.S. Consumer Product Safety Commission, July 1975.

J.J. Persensky and A.M. Ramey Power Lawn Mowers: Evaluation of Anthropometric Foot Probes. National Bureau of Standards NBSIR 77-1294, May, 1977.

V.J. Pezoldt and J.J. Persensky Power Lawn Mowers: Ease of Pull. National Bureau of Standards NBSIR 77-1298, June, 1977.

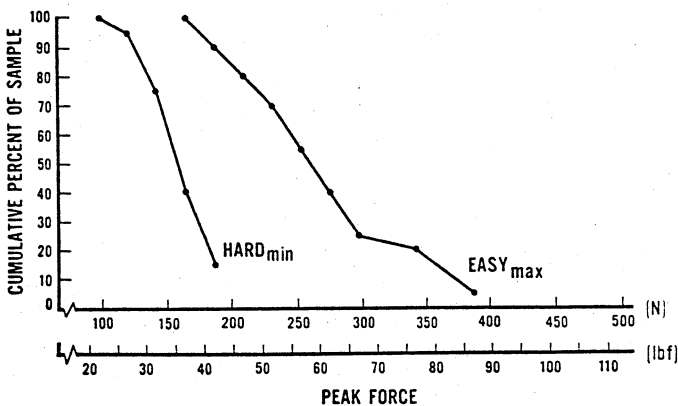
L.G. Porter, C.E. Jones, and J.J. Persensky Some Aspects of Lawn Mower Safety, Status Report to U.S. Consumer Product Safety Commission, December, 1974.

A.M. Ramey and J.J. Persensky Power Lawn Mowers: Time-to-Blade-Access. National Bureau of Standards NBSIR 77-1299, May, 1977.

U.S. Consumer Product Safety Commission. Product Profile: Power Mowers, October, 1976.

U.S. Consumer Product Safety Commission. Power Lawn Mowers, Proposed Safety Standard and Extension of Time. Federal Register, Vol. 42, No. 87 - Thursday, May 5, 1977, 23052-23072.

FIGURE 7
CUMULATIVE DISTRIBUTION OF HARD_{MIN} AND EASY_{MAX}
HANDLE SIMULATOR, FEMALE SUBJECTS



CHILDREN'S CONCEPTIONS OF MEDICINE:
THE ROLE OF ADVERTISING¹

Thomas S. Robertson, John R. Rossiter and Terry C. Gleason
Center for Research on Media & Children, The Wharton School

Abstract

This paper examines three important studies concerning children's orientations toward health and proprietary medicines and the relative impact of advertising on children's beliefs, attitudes, and proprietary medicine usage patterns. The report on the Wharton study is the first data emanating from this recently-completed project.

Overview

An issue of some consequence in the formulation of broadcast regulatory policy is the impact of medicine advertising on children. This issue was heightened in 1975 by a petition to the Federal Communications Commission emanating from the Attorneys General of nineteen states. In their petition the Attorneys General requested a ban on medicine advertising on television between 6:00 a.m. and 9:00 p.m. (Attorneys, 1975).

The substance of this petition was not so much that exposure to medicine advertising leads children to consume medicine, since consumption is presumably mediated by parents, but that exposure to medicine advertising induces receptive attitudes toward medicine-taking. Concern was voiced that such attitudes later become manifest in increased consumption and perhaps over-consumption, or dependency, on these products in adulthood.

The Attorneys General request for a ban on broadcast advertising of proprietary medicines before 9:00 p.m. was subsequently denied by the Federal Communications Commission on the grounds of lack of evidence for the effects suggested:

"... in the absence of empirical evidence to support the claim of a causal connection (between misuse or abuse of drugs and televised advertisements of over-the-counter drugs) it would be unreasonable and arbitrary for (the Commission) to accept the idea that otherwise lawful advertising should be prohibited" (FCC, 1976).

The purpose of this paper is to review three key studies relevant to the issue of advertising's impact on children's conceptions of medicine. In particular, we will focus on some initial results of a study just completed at Wharton, the objective of which is to provide the missing evidence concerning the link between advertising and medicine conceptions. A point of clarification is in order first, however, concerning children's exposure to proprietary medicine advertising. Industry self-regulation codes prohibit the advertising of medicines on "children's programs" (National Association of Broadcasters, 1974). However, these codes apply only to programs "initially designed for children" -- that is, those shows which are concentrated on Saturday mornings. Most children's viewing, however, is to non-children's programs. Among the top fifteen shows viewed by children, only three are shown in time periods covered by the NAB children's codes and 85% of children's viewing is to non-children's programs (Adler, 1977).

¹Funded by National Science Foundation, Grant No. APR76-21130

Research Evidence

There is little systematic research evidence on the effects of medicine advertising on children. Most research concerning advertising and drug products has focused on teenagers, rather than children, and has been concerned with illicit drug use, rather than proprietary medicine use.

The most significant studies to-date are twofold. The first is a study by Milavsky, Pekowsky and Stipp (1975-76) of NBC focusing on television drug advertising and its relationship to proprietary and illicit drug use among teenage boys. The second is a study by Atkin (1975) of Michigan State University, addressing the effects of television advertising on fifth, sixth, and seventh grade children. In this paper we also want to present initial results from the recently-completed Wharton study, which examines the impact of proprietary medicine advertising among third, fifth, and seventh grade children.

The NBC Study

This is a study of teenage boys utilizing a longitudinal panel design. The basic thrust of the research is to relate television viewing to reported usage of proprietary drugs and illicit drugs. The strength of this study is the measurement of television viewing, which is then combined with Broadcast Advertisers Reports, Inc. data which tracks commercials appearing on programs. Thus, the researchers were able to derive a drug advertising exposure measure indicating the total number of advertising seconds for proprietary medicines to which each boy was exposed. This procedure is described in detail by Milavsky, Pekowsky and Stipp (1975-76).

The NBC research findings indicate a positive, but relatively weak relationship between exposure to proprietary advertising and reported usage of proprietary medicines. This relationship is accentuated in homes where there are many proprietary drugs around the house.

By contrast, the NBC research finds a negative relationship between drug advertising exposure & reported usage of illicit drugs -- a finding similar to that of Hulbert (1974) in research with college students. Illicit drugs, however, are not our concern in this paper, nor is there advertising for illicit drugs.

The Atkin (Michigan State) Study

The Atkin study is based on data collected from a sample of 256 fifth, sixth, and seventh grade children selected from schools in urban, suburban and small town areas in Michigan. The basic thrust of the research is the examination of the relationship between television advertising exposure and a relatively rich set of dependent variables concerning children's beliefs, attitudes, and usage patterns for proprietary medicines.

The strength of the Atkin research is in his conceptualization of a set of interesting mediating variables affecting any proposed advertising exposure-medicine usage relationship. The weakness of the Atkin data,

in our opinion, is his measure of "medicine advertising exposure," which was constructed by multiplying amount of viewing to a limited set of television programs by the child's reported frequency of attention to four specific medicine commercials.

Atkin's results may be summarized in terms of the following medicine advertising exposure relationships, all of which are based on sixth-order partial correlations--controlling for grade, sex, social class, scholarship, child's frequency of illness, and parent's approval of medicine use.

- Population incidence of illness. Children with high exposure to medicine advertising perceive that people are more often sick ($r = .14$) and that they more often take medicine (.14).
- Belief in medicine. High medicine advertising exposure correlates with the child's belief in the quickness of relief after taking medicine (.10).
- Illness anxiety. Children with high exposure to medicine advertising worry more about getting sick (.14).
- Approval of medicine. The relationship between medicine advertising exposure and the child's approval of medicine is .12.
- Medicine efficacy. Children with high medicine advertising exposure are more likely to report feeling better after taking medicine (.12).
- Medicine usage. There is a general lack of relationship between medicine advertising exposure and medicine usage (.03).

In general, these results suggest that medicine exposure does, to a certain extent, influence the child's conceptions of illness and medicine. These relationships tend to be accentuated somewhat among the smarter children (as measured by scholastic performance) and among the higher social status children. Other variables such as age and sex of the child, parental attitudes toward medicine, or the child's frequency of illness all show inconsistent patterns.

Neither the NBC nor the Atkin study begins to settle the issue of proprietary medicine advertising's impact on children's medicine conceptions. A major fault characterizes both studies: the reliability of the measures is not reported. Correlations between unreliable measures can be seriously attenuated. Although both studies employed multi-item measures of medicine advertising exposure which are likely to be quite highly reliable, the reliabilities of the "effects" measures are unknown. The unreliability problem is particularly acute in the Atkin study (the only study with children as subjects) where single-item measures were employed. The consequence of unreliability is to produce lower correlations than might have been obtained with reliable measures. The possibility remains that the impact of medicine advertising on children is substantially larger than previously reported.

The Wharton Study

The Wharton Study was conducted in order to present a more conclusive investigation of children's conceptions of medicine in relation to various socialization factors that may be contributory. Exposure to televised medicine advertising was one of these factors.

The sample consisted of 673 children in approximately equal numbers from the third, fifth and seventh grades (ages 8 to 12). The sample ranged from

disadvantaged to upper-middle class families. Approximately equal numbers of boys and girls were interviewed at each grade level. The medicine advertising exposure measure was constructed in the same manner as in the Milavsky et al. NBC study. A random sample of 35 television shows was taken, excluding school hours and late night programming (past 11:30 p.m.). Children were asked to indicate how often they had watched each program on the last four occasions preceding the interview. Broadcast Advertisers Reports were then consulted to determine the number of medicine commercial minutes appearing on each show over the reporting period. The shows were weighted by the medicine advertising count and by viewing frequency, then summed to derive an overall measure of medicine advertising exposure. Reliability for this measure was 73.²

The "effects" variables for this study consisted of: 1) Belief in the efficacy of over-the-counter medicines, 2) Affect toward (liking of) over-the-counter medicines, 3) Intention to take over-the-counter medicines when symptoms occur, 4) Request frequency to parents for over-the-counter medicines, and 5) Usage frequency for over-the-counter medicine. Two additional variables comparable to the Atkin study were also examined -- the child's perception of the population's incidence of illness and an illness anxiety measure.

Each "effects" variable was measured by five self-rated items. These items comprised five categories of over-the-counter medicines: headache remedies, stomach remedies, cold remedies, cough remedies, and vitamin preparations. Reliability coefficients computed over the five items for each variable were: Beliefs .81, Affect .81, Intentions .75, Requests .76, Usage .52, Population Incidence .59, and Illness Anxiety .84.

Results. The findings are shown in Table 1, including the zero-order correlations and the fifth-order partial correlations. The variables controlled for (generally similar to Atkin) are grade, sex, social class, child's illness level, and parental supervision of medicine use.

In general, the magnitude of relationships is similar to Atkin, suggesting a somewhat tenuous association between medicine advertising and the several "effects" variables. The results may be summarized as follows:

- Medicine advertising exposure shows a zero-order correlation of .11 (significant at .01 level) with beliefs in the efficacy of medicine but only .04 at the fifth-order level.
- The relationship builds slightly when medicine advertising exposure is related to affect toward medicine (.14 zero-order and .07 fifth-order partial correlations).
- The relationship improves again when medicine advertising exposure is related to intent to take medicines when ill (.17 zero-order and .10 fifth-order).
- The strongest relationship holds between medicine advertising exposure and reported requests for medicine when ill (.24 zero-order and .18 fifth-order).
- When medicine advertising exposure is related to usage, the zero-order correlation is only .07 and the fifth-order .00.

² Reliability was computed using Cronbach's coefficient α (Cronbach, 1951). This measures the internal consistency of a series of items used to define a variable (see also Rossiter, 1977). The size of α reflects the degree to which the set of items comprising a test is oriented toward a single underlying dimension.

Table 1

Correlations Between Medicine Advertising Exposure and Children's Medicine Conceptions

Variables	Zero-order Correlations	Fifth-order Correlations*
Belief in the efficacy of Medicine	.11 ^b	.04
Affect toward Medicine	.14 ^a	.07 ^c
Intent to Take Medicine	.17 ^a	.10 ^b
Requests for Medicine	.24 ^a	.18 ^a
Medicine Usage	.07 ^c	.00
Population Incidence of Illness	.00	-.10 ^b
Illness Anxiety	.22 ^a	.13 ^a

* Controls for grade, sex, social class, illness level, and parental supervision of medicine-taking.

a Significant at .001 level

b Significant at .01 level

c Significant at .05 level

- The relationship between advertising and the child's perception of the population incidence of illness is .00 (zero-order) and -.10 (fifth-order).

- Finally, the medicine advertising exposure/illness anxiety relationship is .22 (zero-order) and .13 (fifth-order).

Interestingly, there is an absence of an advertising/usage relationship. However, requests for medicines to parents represents the critical response measure for children's consumer behavior. This is because parents execute most purchasing for children (especially for medicines) and requests, therefore, should be regarded as the "response" over which the child has control. Analysis of the advertising exposure/requests relationship indicates that it is moderate (.18 fifth-order) and highly significant in a statistical sense (.001 level).

Conclusions. What have we learned at this early stage of our analysis? Generally, television advertising of proprietary medicines has a rather moderate effect on children's beliefs, affect, intentions and request behavior. Interestingly, the effect increases as we move from beliefs to requests (fifth-order partials of .04 versus .18). There is no obvious impact of television advertising on actual usage of medicines. Apparently, usage behavior is a function of other factors, most particularly illness level and parental mediation.

In order to understand children's conceptions of medicines, television advertising as an explanatory factor must be placed in the context of a multi-mediational model of effects. The examination of such an enriched model will comprise the next phase of our analysis. In particular, we wish to understand the family mediation process and the role of multiple information sources in affecting children's conceptions of medicine.

References

- Richard Adler, et al., The Effects of Television Advertising on Children. A report submitted to the National Science Foundation, April, 1977.
- Charles K. Atkin, The Effects of Television Advertising on Children: Survey of Pre-Adolescent's Responses to Television Commercials. A report submitted to the Office of Child Development, H.E.W., July, 1975.
- Attorneys General, "Petition of the Attorneys General to Promulgate a Rule Restricting the Advertising of Over-The-Counter Drugs." Before the Federal Communications Commission, Washington, D.C., August, 1975.
- Broadcast Advertisers Reports, Inc., "Network TV: Brand Product and Parent Company Schedule Detail and Expenditure Estimates for the Weeks Ending February 20,27, March 6,13, 1977." New York: BAR, 1977.
- L. J. Cronbach, "Coefficient Alpha and the Internal Structure of Tests," Psychometrika, Vol 16, September, 1951, 297-334.
- Federal Communications Commission, "News Report No. 14747," December 10, 1976-B.
- James Hulbert, "Applying Buyer Behavior Analysis to Social Problems: The Case of Drug Use," in Proceedings of the American Marketing Association. ed. Ronald C. Curhan, 1974, 289-292.
- J. Ronald Milavsky, Berton Pekowsky, and Horst Stipp, "TV Drug Advertising and Proprietary and Illicit Drug Use Among Teenage Boys," Public Opinion Quarterly, Volume 39, Winter 1975-76, 457-481.
- NAB, "Children's TV Advertising Statement of Principles," Code News, The National Association of Broadcasters, Volume 7, No. 6, June 1974.
- John R. Rossiter, "Reliability of a Short Test Measuring Children's Attitudes Toward TV Commercials," Journal of Consumer Research, Volume 3, March, 1977, 179-184.

CONSUMER BEHAVIOR AND NUTRITION: PREVENTIVE HEALTH PERSPECTIVES

M. Venkatesan, University of Oregon

Abstract

There is increasing interest in consumer behavior and nutrition. However, the present attempts are inadequate inasmuch as they follow the "curative" approach. What is needed is a "preventive health approach." Such an approach recognizes that changing consumer nutritional habits is a long-term phenomenon. A brief summary of the current state of our knowledge in this area is provided and suggestions for future research are outlined.

Ever since the Federal Trade Commission issued a proposed Trade Regulation Rule on food advertising aimed at regulating the claim and information contained in the advertisements for food products, considerable attention is being devoted to consumers and nutrition. Much of this attention has been directed in the forms of suggestions -- some of which resulted from recent studies -- towards the formats in which such information should be provided [Kendall, 1977; Scammon, 1976], the type of information that should be provided [Bettman, 1975], and the attributes of such information [Quelch, 1977]. Other studies have concentrated on the current level of knowledge and the use of such knowledge by consumers [Jacoby, 1976, USDHEW, 1974, 1975].

It is interesting to note that these flurries of activities did not deal with the reasons for the attempts of the Federal Trade Commission and others to provide nutritional information. The current attempts at disclosure are based on the belief that there is nutritional deficiency among consumers and is based on the findings of three major studies, viz., The Ten State Nutrition Survey [1972], First Health and Nutrition Examination Survey (HANES) [1974] and the USDA Survey of Dietary Levels of Households in the United States [1965]. There are unresolved controversies surrounding the interpretations of the findings from these three studies and the segments of population to which such findings properly relate. Irrespective of these difficulties, the philosophy behind the TRR of the Federal Trade Commission, and therefore, of the studies that have attempted to research and suggest ways in which to communicate the nutritional information have all taken a "curative" approach to the problem. That is, a "malnutrition" problem is perceived to exist among consumers and therefore the researchers are suggesting "cures" to the policy makers. It is clear that our studies of consumer knowledge of nutrition [Jacoby, 1976] and their use of such information and our suggestions [Bettman, 1975, Scammon, 1976, Venkatesan, 1977] have all been influenced by this "curative" perspective of the problem of nutritional deficiency among consumers. In short, we have taken the same approach as much of the medical practitioners and which has come under increasing criticism. The reason for this can be traced to lack of clarity in our thinking with regard to nutrition and consumer behavior and our haste to apply our "consumer behavior approach" to this area of human behavior and assuming the results will be similar to buying behavior research for a variety of products and services.

We have finally come to understand that in order to affect the behavior of consumers in the area of nutrition, we have to be concerned not with the "curing" aspects of nutritional deficiency, if any, but with the

preventive health behavior aspects of consumer behavior. It is becoming increasingly clear that more and more consumers are exhibiting concerns about food additives, ingredients of food products, trace minerals, etc. reflecting an interest in prevention. We, as a society, are becoming more convinced that prevention is both simpler and cheaper than cure in all aspects of our behaviors. There is increasing emphasis on prevention of every kind -- medical, occupational, environmental, and nutritional. For purposes of this paper, preventive health behavior for consumers in the nutritional area can be defined (with a slight modification of the existing definition of health behavior of Kasl and Cobb, 1966) "as any activity undertaken by a consumer who believes himself/herself to be healthy, for the purpose of preventing nutritional inadequacies.

In general, there are two types of prevention activities that can be undertaken by individuals, viz., passive prevention and active prevention. Passive prevention does not involve any active participation by the consumers to get prevention. An example of such prevention might include fortification of all food products with required levels of vitamins, minerals, etc., or elimination of all additives/ingredients (and even products) which are not fully tested and which are suspected or even remotely linked to some cancer producing agent. As is obvious, passive prevention measures will not, for the most part, involve the consumer. However, such passive preventive measures are not feasible because providing fortification of food products becomes prohibitively expensive. Secondly, much is not known about the long term effects of such fortification efforts. Thirdly, not all of the individuals require the same level of nutrition intake.

Much of our problem in the nutritional area (deficiency, obesity, etc.) are "self-imposed" such as excessive eating, eating of junk foods and improper diets, etc. Nutritional knowledge and use of such knowledge in the choice and consumption of nutritionally beneficial food products fall under "consumer behavior," as such activities involve consumer awareness about nutrition and relating such knowledge to foods and food products and consumption of such products relate to consumer choice processes. Here, our understanding of consumer behavior and applying it to nutritional settings might be of help to both the producers of these products and to the policy makers. Thus our concern is with consumers' active prevention behavior and that implies that the individuals have specific things to do, some of which might affect and/or change their life styles. That is, consumers are fully expected to actively participate to gain benefits of prevention. Even if such preventive measures are taken by consumers, there is no guarantee that any illness or deficiency will be prevented for the individual nor can it be claimed that such prevention may lead to elimination of any disease, etc.

The active prevention notion in the nutritional area involves social change concepts, that is, changes in the life styles of individuals over a long period of time and deliberate intervention programs may become part of change strategies. Some claims have been made that such social changes can be brought about by "social marketing." [Kotler and Zaltman, 1971]. We do not currently have ample demonstrations or empirical

evidence of success in this area. There are also a number of ethical and public policy considerations in any attempt to implement active prevention measures by consumers in the nutritional area. There are still controversies regarding the RDA levels that are set as guidelines and there will be problems in policy makers attempting to suggest not only consumption of specific foods (eggs vs. beef, etc.) but consumption of specific brands of food products.

Current State of Our Knowledge

The Food and Nutrition Surveys (1974, 1975) of the Food and Drug Administration are the most comprehensive studies to date. By and large, these two surveys indicate that a substantial proportion of consumers are knowledgeable about the concept of "well-balanced diet" and only a small proportion of consumers have indicated that they were not getting a well-balanced diet and this was not due to any lack of awareness on their part, but that it was due to other reasons such as "fussy eating" and too much time spent away from home and the like. The Red Book surveys (1974, 1976) also indicated that a very high proportion of their respondents are aware of the basic food groups, and recognized the need for well balanced meals and the like.

If the awareness level is high, why then is their consumption behavior presumably not consistent with their awareness? For one thing, none of the national surveys have found any serious concern or perception of a serious problem with nutritional deficiency. Most consumers are concerned with obesity and other similar problems. Bauman [1973] correctly identified the reasons for lack of concern when his surveys found that most housewives believed that they were serving their families with a variety of basic foods over a period of time, which they believed will achieve nutritional balance. The second reason Bauman attributed for nutritionally inadequate eating habits is the breakdown of traditional three-meals-a-day pattern of eating behavior and the growing tendencies of families to eat more meals away from home. Another reason advanced by some is that while there may be general awareness of nutrition, the need for well-balanced diet and the like, consumers generally lack information on specific nutrients. Their problem is compounded by the fact that consumers have difficulties connecting their already meager knowledge of specific nutrients with the food products in general and with specific brands of food products in particular. There may also be a confusion factor, as Lachance [1973] observed:

The consumer wants a balanced diet, but the confusion over how to attain it is fantastic. Not only are there more foods to choose from, but we nutritionists have contributed to the confusion. One can take any one of several college level textbooks on nutrition to learn that the good sources of protein, several vitamins, and minerals are organ meats (e.g., liver), egg, milk, cheese, etc., only to advise in a separate chapter or article that to avoid heart disease one must limit the intake of -- you guessed it -- organ meats, milk, eggs, certain cheeses, etc.

Finally, as the surveys have repeatedly found that the unshakeable belief among consumers seems to be that one can get enough nutrition from eating a variety of foods from the supermarket and this contributes to the problem of consumer education. There are other reasons and there are a number of studies that point out problems and difficulties in attempting to change the eating habits of consumers. [For a summary of these studies and a critique, see Kendall, 1977].

Proposal for Research

Research in how consumers make choices of food products in the supermarket is very sketchy and much less is known about how they combine their nutritional information with actual choice of branded products. It is obvious that food choices are not the same as brand choices of food products. The recent research studies that deal with consumers evaluative criteria for food products and with attributes of food products and the type of information that needs to be provided are all concerned with way "ought" to be provided and not with how and why consumers make their choices of food products the way they do. For example, we know very little about the criteria consumers seem to use in choosing brands of food products. Bayton [1968] has reported that he had found from his research in the purchase of food products that consumers had the following parameters, which he grouped into seven categories:

I. Nutrition Parameters

1. Body growth needs
2. General health needs (rather than specific health needs)
3. Vitality; energy
4. Energy "carry-through" (concern over longlasting energy)

II. Economic parameters

5. Price, per se
6. Value (what you get from the money)

III. Sensory--aesthetic Parameters

7. Taste-aroma-appearance complex
8. Refreshment (especially "coolness" relief of thirst)

IV. Personableness Parameters

9. Personableness-in-general (lively; good complexion, bright and sparkling eyes; general attractiveness)
10. Sex personableness
Males--vigorous, athletic; masculinity
Females--lovely complexion; nice figure, femininity

V. Appropriateness (suitability to my kind of person or to given situations)

11. Religious-cultural appropriateness
12. Age--group appropriateness (milk for children, coffee for older ages)
13. Status-group appropriateness (class-relatedness of some foods)
14. Social setting appropriateness (family privacy; intimate friends, special guests; restaurants)

VI. Convenience

15. Convenience in purchasing; availability
16. Convenience in storing
17. Convenience in preparation
18. Convenience in serving
19. Convenience in consumption

VII. Health apprehensions

20. Weight apprehension
21. Cardiac apprehension
22. Contamination apprehension (pesticides, bacterial, animal medications, atomic fall out)
23. Allergies

There is not much research evidence available on these parameters. However, it is clear that jumping to conclusions or extrapolations from findings on the use of nutrient labeling to the whole area of nutrition and consumer behavior is both imprudent and premature. Secondly, changing or influencing consumer behavior in this area, as pointed out earlier, is a long-term problem. In a study [Wilson, 1972] attempts were made to educate six families, on a longitudinal basis, on nutrition information. The resultant improvements in their food habits was minimal. Thus, quick solutions and grandiose short-term suggestions are not likely to succeed in changing the behavior of consumers. What is needed is a program of painstaking experimental research for a long period and implementation of the research based approach by policy makers will in all likelihood increase the chance of maximum benefit at least cost for the society and enhance the nutritional well being of our society.

References

- Howard E. Bauman, "What Does the Consumer Know About Nutrition?" Journal of the American Medical Association, Vol. 225, No. 1, July 2, 1973, pp. 61-62.
- James A. Bayton, "Psychodynamics of Food Consumption." Paper presented at the 12th Annual Nutrition Forum, Virginia Council on Health and Medical Care. Richmond, Virginia, April 22, 1968.
- James R. Bettman, "Issues in Designing Consumer Information Environments," Journal of Consumer Research, Vol. 2, December 1975, 169-177.
- First Health and Nutrition Examination Survey, United States, 1971-72. (Preliminary Findings) Dietary Intake and Biochemical Findings U.S. Department of Health, Education, and Welfare. Public Health Service, Health Resources Administration, National Center for Health Statistics, Rockville, Maryland, Jan. 1974.
- Jacob Jacoby, R. W. Chestnut, and William Silberman, "Consumer Use and Comprehension of Nutrition Information." Paper No. 163 Purdue Papers in Consumer Psychology, 1976.
- Stanislav V. Kasl, and Sidney Cobb, "Health Behavior, Illness Behavior, and Sick Role Behavior," Archives of Env. Health Vol. 12 (February 1966) pp. 216-267.
- Kenneth W. Kendall, Methods of Presenting Nutrient Information in the Media. An Experimental Investigation. Unpublished Doctoral Dissertation, University of Iowa, 1977.
- Phillip Kotler, and Gerald Zaltman, "Social Marketing: An Approach to Planned Social Change," Journal of Marketing, July 1971, pp. 3-12.
- Paul A. Lachance, "Nutrition and the Public Health," Presentation before the New York Section of the American Association of Cereal Chemists, New York City, April 3, 1973.
- Red Book Reader Survey. How Do You Feel About Food? What Do You Know About Food? The Results of a Red Book Reader Survey, May, 1974.
- Red Book Nutrition Study. Princeton, New Jersey. The Gallop Organization, 1976.
- Debra Lynn Scammon, The Effects on Consumers of Varying the Amount and Format of Purchase-Relevant Information. Unpublished Doctoral Dissertation, University of California, Los Angeles, 1976.
- Ten State Nutrition Survey, 1968-70. U.S. Department of Health, Education, and Welfare. Health Services and Mental Health Administration, Center for Disease Control, Atlanta, Georgia, 1972.
- United States Department of Agriculture. Household Food Consumption Survey 1965-66. Report No. 6 Dietary Levels of Households in the United States, Spring 1965. (Agricultural Research Service)
- U.S. Department of Health, Education and Welfare. Public Health Service, Food and Drug Administration. Consumer Nutrition Knowledge Survey. Report I 1973-74. (DHEW Publication No. FDA 76-2058)
- Consumer Nutrition Knowledge Survey Part II, 1975. DHEW Publication No. FDA 76-2059.
- M. Venkatesan, "Preventive Health Care and Marketing: Positive Aspects." A paper presented to the Preventive Health Workshop, American Marketing Association, April, 1977 (in press).
- Consumer Nutrition Knowledge Survey, "Providing Nutritional Information to Consumers." Paper presented at the special NSF/MIT Conference. "Consumer Research for Consumer Policy," Cambridge, Mass., July, 1977 (in press).

ENVIRONMENTAL PROTECTION FOR THE NON-SMOKER:
CONSUMER BEHAVIOR ASPECTS OF ENCOURAGING NON-SMOKING

Lawrence H. Wortzel, Boston University
Roberta Clarke, Boston University

Abstract

It is becoming more widely recognized that cigarette smoke is harmful to non-smokers, yet there has been very little study of the ways in which non-smokers can protect their environments from smoke. This paper reviews and recasts studies on cigarette smoking to bring their results to bear on the problem of non-smoker environmental protection. It also offers some ideas for both research and action.

Introduction

The existing research and related literature on cigarette smoking can be broken down into three major categories. The first category, which makes up the bulk of the literature, details and seeks to prove the physical and health-related damage experienced by those individuals (and experimental animals) who smoke cigarettes. The second category of literature, somewhat less substantial in nature, directly addresses the smoke and those who would wish to influence the smoker not to smoke. This literature is concerned with the smoker's own good.

It is the third category, however, which this paper addresses. In this category is literature which seeks to educate the non-smoker about how to prevent the smoker from lighting up a cigarette in the non-smoker's presence for the non-smoker's good. There are, at present, few studies and little literature dealing with this topic. However, we believe that this category will soon become more actively studied for several reasons.

First, while it has long been widely accepted (at least on the cognitive if not the affective level) that cigarette smoking is harmful to the health of the smoker, only recently has it been recognized that smoking can also be deleterious to the health of the non-smoker who is in close proximity to smokers. Scientific evidence that cigarette (as well as pipe and cigar) smoke is harmful to the non-smokers who are exposed to it (U.S. P.H.S., 1972) points to such harmful effects as eye and nose irritation which is immediately apparent but transitory; children particularly are sensitive to these effects (Lugnette *et al.*, 1970). Of a more lasting nature, however, is the possible increased probability of lung cancer and emphysema, neither of which is apparent to the non-smoker at the time of exposure. In fact, many non-smokers, unless they are specifically allergic to cigarette smoke, have probably been relatively unaware of the long term hazards of smoke from others' cigarettes.

However, this situation can be expected to change. There are significant efforts being made to inform non-smokers of both the short and long term hazards of exposure to smoke. The American Lung Association, for example, publishes a pamphlet, Second Hand Smoke, the majority of which is devoted to explaining the dangers of smoke. This pamphlet also briefly suggests some defensive actions which non-smokers can take to protect their environments. Moreover, Action on Smoking and Health (ASH) offers bumper stickers, buttons and small message stickers which either promote non-smoking or request that the reader not smoke. GASP (Group Against Smokers'

Pollution) of Massachusetts, another consumer activist anti-smoking, exists to publicize the hazards of second-hand smoke and to support legislation aimed at insuring non-smokers' rights.

The ultimate solution to the second hand smoke problem may be the enactment and vigorous enforcement of anti-smoking legislation. Indeed, significant progress has been made in this direction, with the establishment of "no smoking" sections in airplanes and (some) restaurants as well as "no smoking" regulations in some public places. But, in the short term, the smoker must still defend against smoke in many of the environments in which he or she finds himself or herself.

A second and related reason why we expect this area (the protection of the non-smoker's air rights) to become more heavily researched is that the growing emphasis on preventive health care has made us all more aware of personal health risks and more active in seeking to maintain our health. Why jog, for example, to fill our lungs with clean air and then breathe others' smoke?

As noted previously, whatever smoking-related health education literature there was existed primarily to educate the smoker to the dangers of smoking. This literature is essentially passivist with regard to whatever health risks smoking presented to the non-smoker. However, preventive health care education is now beginning to extend itself to educate the non-smoker in smoking-related preventive care; that is, how to prevent smokers from smoking near you. The growing activist stance in non-smoker preventive care is certain to have an effect upon future research in this area.

While there is not much literature directly on the topic of non-smoker environmental protection, there is a considerable body of literature, some of which has been reviewed, (Fishbein, 1977) which can be considered as having some relationship to this anti-smoking problem; however, the literature has never been examined from the non-smokers point of view. If the literature has been reviewed in any kind of action format, that format has been one aimed toward either encouraging non-smokers not to start or encouraging smokers to stop (permanently). The question here is a much more temporary one; it is concerned only with encouraging non-smoking for short or finite lengths of time in specific places. Thus, the primary emphasis is on identifying consumer behavioral reasons that are amenable to immediate attack.

This paper, therefore, deals with the short-run (non-legislative) problem of activist non-smoker environmental defense: how the non-smoker can defend whatever environment he or she is in and only for the length of time he or she happens to be in that environment. The paper relies upon consumer behavior analysis as a basis for selecting marketing tools and promotional context which can be effectively used to inhibit smokers in distinct situations.

The studies which have been reviewed fall into two broad areas: 1) internal cues: smoking and personality/social behavior characteristics, and 2) external cues: factors which trigger the smoker to light up a cigarette. Both

areas will be discussed; included in the discussion are suggestions for strategies and tactics that might inhibit short term smoking. The format will be first to review the studies and then to suggest action strategies. This review will be in no sense exhaustive as the authors' own reasonably exhaustive review indicated a considerable degree of redundancy in the literature. Therefore, only a representative group of studies deemed to have actionably useful findings will be reported here.

In addition to the literature on psychosocial correlates of smoking, there is also a literature on physiological correlates. Some of this literature (Schachter, 1977) presents evidence that cigarette smoking produces physical dependency if not addiction. This literature will not be addressed here; while a recognition of the role that physiology plays may be significant in inducing an individual to stop smoking permanently, there is not any strong evidence that physical symptoms are so intense as to prevent smokers from postponing a smoke for a finite length of time. The review begins, therefore, with internal cues.

Internal Cues: Personality and Social Behavior Characteristics

There have been quite a few studies covering the relationship between smoking and personality and various social behavior characteristics. None of these studies have demonstrated that any of the factors that were studied actually caused smoking; however, their presence may offer some clues to the non-smoker who wishes to take action.

Review of Studies

A number of studies examine the motivations for and correlates of smoking among adolescents, since many smokers begin smoking during adolescence.

Adolescents begin their smoking in a peer-group atmosphere which they perceive as accepting of smoking. In a national study including both smokers and non-smokers conducted in 1975 by Yankelovich, Skelly and White, over 80% of teenaged boys and girls said they usually think of teen-agers as smokers, and 67% said they usually think of teachers as smokers. This indicates that teen-agers believe smoking was an activity accepted by both their peers and their school authority figures. (U.S. Dept. of HEW, 1977). Perhaps the most insightful description of adolescents' motivations for smoking is offered by Neeman & Neeman, who comment:

It is possible that the act of smoking itself is, in part, a component of the communication process, namely an "emblem" or display: the blowing out of a plume of smoke has been likened to the display of a peacock's plumes, decoded as an emblem of status or dominance vis-a-vis peers of the same or opposite sex, communicating the smoker's toughness, independence, and a precocious type of sexual maturity, and conveying social acceptance and recognition. Smoking also provides a means of offering or accepting social contact. (Neeman & Neeman, 1975)

These comments are well supported and sex-differentiated by data from the survey conducted by Yankelovich *et al.*, which was referred to earlier. This survey concluded that cigarette smoking among teenaged boys was associated with social uneasiness, the desire to be popular with girls and demonstrating one's masculinity. Girls, on the other hand, do not see smoking as a social asset, but rather as a sign of a rebellious independence, and a somewhat defiant social maturity. For example, a much larger proportion of teenaged girl smokers, as compared to non-smokers, said they drink alcohol and had had

sexual relations (U.S. Dept. of HEW, 1977).

Studies have also been conducted among college students. Some of these studies have used general personality inventories to try and differentiate between smokers and non-smokers. Schubert gave the MMPI, along with measures of tobacco usage to two samples of college students, totalling 1270 respondents. His results were not consistent between the two samples when these results were analyzed in terms of traditional MMPI categories. There was no consistent personality differentiation between smokers and non-smokers. However, an item analysis did indicate some constant differences. Smokers' self-descriptions indicated they were more likely to be bored, impulsive and thrill-seeking, to behave in a socially unacceptable fashion, to ascribe this trait to others and to have masculine traits. These characteristics were described generally as arousal-seeking (Schubert, 1965).

These findings were both reinforced and extended by a later study using 517 male college students as subjects and the Boston University Personality Inventory as the instrument. Heavy smokers (as compared to non-smokers) were more impulsive, defiant and danger-seeking. These results also were reported as suggesting that smokers used smoking as a means of warding off manifest distress conditions such as tension, irritability and boredom (Jacobs & Spilken, 1971).

There is also some evidence that adult smokers have considered their behavior as deviant but not defiantly so. As early as 1966, 45% of all smokers (then currently smoking) interviewed in a National Survey agreed that "cigarettes are morally wrong." (Nuehring & Markle, 1974). During the years between 1966 and 1974, the authors argue, many smokers have become 'repentant deviants' (as opposed to the younger defiant deviant) who admit the reprehensibility of their actions and by doing so gain license to persist in smoking. But the more mature smoker has integrated smoking into his or her personality and/or into his or her psychomotor behavior. When one becomes an established, regular smoker, smoking becomes a means of telling others about oneself, but the acts associated with smoking become habitual and autonomic. Possibly, in fact, smoking becomes an idiosyncratic behavior akin to foot tapping, pencil twirling, nail biting and the like (Dunn, 1973).

Action Implications

These findings present some intriguing yet contradictory possibilities for action. However, if non-smokers are to take these actions, they would have to do so as a group as the action possibilities might best be defined as media advertising in public places. The content of this advertising could be communicated by individual non-smokers as well, but less successful results would be expected.

The opportunities for action in terms of "advertising" would differ between more experienced vs. newer smokers, especially if the newer smokers and teen-agers or college students. Among more experienced smokers, the general tenor of messages should be (1) to emphasize that smoking in that particular place is deviant behavior and that the only way to repent is not to smoke and (2) to raise consciousness that one is smoking in that place if one does so, since smoking for some individuals is so habitual that they are at a conscious level momentarily unaware that they are smoking.

On the other hand, messages such as these should cause young, newer smokers, especially males, to light up a cigarette. These messages would reinforce deviant (smoking) behavior among this group since teen-aged male smokers are defiant deviates; moreover, raising consciousness about the act of smoking raises the impor-

tance of the situation as an opportunity to display one's masculinity or one's social maturity as the case may be. Therefore, it should be more effective instead for advertisements to portray smoking in such environments as unmasculine, socially immature and a sign of social weakness. Unfortunately, this type of advertising (which would probably be communicated in print through buttons, brochures, posters and youth newspapers, given the higher cost of broadcast media) is not likely to be very effective principally because of the importance of the peer group in guiding smoking behavior.

In summary, the studies of personality and social behavior characteristics suggest a possible line of attack for producing non-smoking behavior among experienced mature smokers but a considerably less attractive line of attack for younger, less experienced smokers. Since specific attitudes and behavior are usually better predictors of action than the generalities discussed in the above studies, it is possible that in the future more specific research could direct us toward more effective action.

External Cues: Factors that Trigger Smoking among Smokers

There have been several studies focusing on relationships between smoking and various external cues. First, we will examine those studies with indirect action implications, followed by studies that indicate action directly. This second section will also rely to some extent on non-academic literature and information provided by anti-smoking groups, as such material considerably broadens the action possibilities open to the non-smoker.

Review of Studies-with Indirect Action Implications

Shapiro *et al.*, studied 750 smokers to determine what links, if any, existed between environmental events and "lighting up." They found that it was usual for smokers to smoke on specific occasions although these occasions varied across people. There were, however, some situations that were popular among all smokers. These were situations that were stressful, required concentration, or were relaxing (Shapiro, *et al.*, 1977). In these situations, individual smokers originally began to smoke to reduce negative affect or to cope with an unpleasant or stressful situation. For others, smoking was initiated to experience or to increase the experience of positive affect, excitement, relaxation and enjoyment associated with the situation. Over time, however, what is experienced by the individual in many such situations is neither enhanced nor dampened by smoking, but instead the specific situation becomes a strong cue to smoke and the individual does so without becoming aware that he or she is doing so (Tomkins, 1966).

Herman, using eating behavior as an analogy, divided smoking cues into external (situational, such as other people smoking) and internal (deprivation of cigarettes). He conducted an experiment using approximately 100 college students as subjects in which internal and external cues were manipulated for both light and heavy smokers. He found that heavy smokers were affected primarily by internal cues (although they were not insensitive to external cues), while light smokers were affected by both internal and external cues (Herman, 1974). It would seem fruitful to explore further the relationship between smoking and external cues since these are "marker-controllable" variables. There is further evidence that external cues are important in the situations with which we are concerned.

The nature of external cue situations may differ across social classes. Meyer *et al.* point out that many blue collar jobs, for example, prescribe or proscribe the amount of smoking that can take place during work. White

collar workers, on the other hand, are usually free to smoke on the job (Meyer *et al.*, 1973). Thus, blue collar workers may be exposed to many fewer external cues than are white-collar workers and therefore may be differentially (either more or less) sensitive to external cues. Thus, Herman might have had different results had he used blue collar workers instead of college students.

There has been some attempt to identify some of the specific situations providing external cues that trigger smoking. Hochbaum, early on, pointed out that smokers link smoking with entering a room where there are people or with seeing someone else smoke (Hochbaum, 1965). Foss reports on a study of 87 present and former smokers; both groups smoked primarily in social situations or social gatherings; 54% of present smokers and 79% of former smokers reported having done so (Foss, 1973). Both Hochbaum and Foss, however, say that the overt act is simply a manifestation or expression of some less apparent but deeper feeling such as anxiety or inadequacy.

In summary, while both internal and external cues influence smokers' behavior, much smoking is done in social situations in which external cues are present in abundance. These cues may serve to trigger smoking, for example, by reminding the smoker of the anxiety inherent in the situation. There is also evidence that people continue to smoke in situations that no longer produce affect, thus suggesting that cues can simply produce smoking behavior for no other reason than the presence of the cue.

Action Implications

With respect to external cues, essentially there are two smoker populations to be considered: those for whom smoking in a given situation produces affect and those for whom the situation is merely a cue, and no affect is produced. The action implications are different in each situation. Where affect is produced, the goal of action is to modify the affect, providing substitute sources of gratification or relieving anxiety as indicated. Where there is no affect present, it would seem that the goal of action is simply to make the smoker aware that he or she is smoking for no particular reason.

The action problem is more difficult when the smoker is motivated primarily by affective cues. Perhaps, the best that can be done is for the non-smoker to recognize the motivation and simply ask for delay.

Review of Studies that Indicate Action Directly

This section of the paper departs slightly from the previous format in that action ramifications will be discussed with the studies or will be suggested independent of experimental data. Non-verbal action possibilities will be discussed first followed by direct verbal communication strategies.

Two relevant studies deal with behavior under the presence or absence of specific external cues, essentially suggesting non-verbal active strategies. These two will be reviewed first.

One such study found that women (but not men) were more likely to smoke when in proximity to others who were smoking (Tryon *et al.*, 1977). The other concerned college students' smoking in a classroom without ash trays. Students refrained from smoking either because of fear of criticism from the instructor or because of guilt about dropping ashes on the floor (Ritter & Holmes, 1969).

These two studies suggest some specific actions. The first study suggests that the non-smoker probably needs

to keep as much territory as possible around him or her smoker free; if a woman is more likely to smoke when in proximity to other smokers, she may want to do so when juxtaposed between the non-smoker and a group of smokers. Obviously, this also suggests segregating all smokers from the non-smoker rather than intermingling smokers and non-smokers. The second study, besides indicating the obvious strategy of ash tray removal, argues that in the non-smoker's own environment, placing an expensive-looking rug where visitors sit could silently inhibit others from smoking (but at some risk).

Commercial programs which help smokers to stop smoking also want to protect their successful graduates from having to deal with smokers. Their recommendations for non-verbal action included not only the removal of ash-trays and matches, but also the wearing of anti-smoking buttons (which they point out may label the button wearer as a fanatic or "nudnik" but are also to some extent successful in temporarily postponing smoking behavior). The balance of the studies to be reviewed deal with specific communication content of various types, each of which suggests some possibility for action.

Horn & Waingrow have proposed using mastery as a motivation to get people to stop smoking. They argue that, for some people, inability to keep control of themselves is more frightening than the health hazards attached to smoking. Thus, presenting smoking as out-of-control behavior may motivate some smokers to give it up (Horn & Waingrow, 1966). Such a strategy might be dubious over the long-term, but for short periods of time could be an extremely effective bit of personal selling. The non-smoker could ask the smoker, "are you so out-of-control that you have to smoke right now?"

Another stop-smoking motivation might be humiliation. Premack posits that it's acutely humiliating to be identified as belonging to an interdicted class (Premack, 1970). To the extent that smokers view themselves as deviants and are really contrite about their position, it can be speculated a humiliating message might work, "Gee, I didn't think anyone in your position would smoke smoke...". However, the success of such a message might depend also on the relative status of the smoker and the non-smoker.

Anti-smoking consumer groups and smoke-ending programs have developed a grocery list of actions which non-smokers might take to protect themselves from second hand smoke. However, some of these groups take pains to point out that non-smokers must first believe they have the right to take these actions against smokers before they will do so. Therefore, the groups believe the first step is to educate the non-smoker as to his or her rights. This has been done by (1) indicating to the non-smoker that smokers are now a minority (46%) of the U.S. population, in a country of majority rule and (2) educating the non-smoker to the dangers he or she may be inflicting on him or herself by breathing second hand smoke. Second hand smoke contains such poisonous substances as ammonia, benzene, carbon monoxide and hydrogen cyanide. The anti-smoking groups' contention is that if the non-smoker recognizes the risks of breathing second hand smoke, he or she is more likely to take an activist stance in attempting to prevent smoking in her or his immediate area.

These groups have also raised the question of whether non-smokers as a group should also attempt to educate smokers about the dangers of second hand smoke. All communications from non-smoker to smoker all likely to be more effective if the smoker is able to appreciate the physical and psychological irritation which smoke causes to the non-smoker.

The communications suggested by anti-smoking groups which a non-smoker may use to temporarily prevent nearby smoking tend to fall into three categories: courtesy,

empathy and medical. Courtesy communications such as "Please don't smoke," or "Could you wait until we are off the elevator before you smoke" are viewed as sometimes effective but also sometimes hostility-arousing as the smoker may believe this his/her rights are being trodden on.

The empathy communication, which tends to be a good deal more personal, is best used by a previous smoker who has stopped smoking permanently (or by a non-smoker who is willing to lie). The message is basically "Please don't smoke right now. I am a former smoker and being so close to someone who is smoking makes me extremely uncomfortable (makes me feel like smoking, makes me anxious, etc.) If you could just wait until I'm gone..."

However, the literature of the anti-smoking groups seems to favor the medical communication which clearly indicates a negative physical effect suffered by the non-smoker due to the smoker's smoke. Complaints of eye and nose irritation or of general allergic reactions to smoke may produce the desired results. However, the stronger the medical complaint, the greater the likelihood of non-smoking behavior, according to the anti-smoking groups. Thus, complaints of emphysema and asthma by the non-smoker to the smoker appear to be the most effective in postponing smoking behavior.

Finally, if all else fails, one may look to Powell and Axrin. They studied the effect of electric shock on cigarette smoking; they found that the effects of shock worked in the short run, as the number of cigarettes a smoker in this experiment smoked were a function of the intensity of the electric shocks he received (Powell & Axrin, 1968). Perhaps, as a last resort, timid non-smokers should install electric chairs in their offices.

Conclusions

Modifying smokers' behavior, even for short periods of time, is a rather complex process and one that should be tailored both to the situation and to the smoker. It appears that group action on the part of non-smokers is likely to be effective only in the short-run. Soon, for example, if individual non-smoker action increases, smokers might start carrying their own ashtrays; moreover, their willingness to believe that the coughing non-smoker next to them has emphysema will decline as the number of self-proclaimed emphysemics increases.

In any event, there are significant opportunities to do research on the subject.

References

- W.L. Dunn, "Methods and Models Applied to Motivation in Cigarette Smoking," in Dunn, W.L., ed., Smoking Behavior: Motives and Incentives, V.H. Winston & Sons, Washington, D.C., 1973.
- Martin Fishbein, Consumer Beliefs and Behavior with Respect to Cigarette Smoking: A Critical Analysis of the Public Literature, unpublished report prepared for the staff of the Federal Trade Commission, May 1977.
- Robert Foss, "Personality, Social Influences and Cigarette Smoking," Journal of Health and Social Behavior, 14 (September 1973), 279-286.
- C. Peter Herman, "External and Internal Cues as Determinants of the Smoking Behavior of Light and Heavy Smokers," Journal of Personality and Social Psychology, 1974, Vol. 30, No. 5, pp. 664-672.

Thomas S. Robertson, The Wharton School
and Lawrence H. Wortzel, Boston University

Abstract

Mass media have considerable potential for affecting health behavior. The pervasiveness of mass media and the exposure levels of broad segments of society suggest that mass media may be an important information source regarding health and a relevant socialization force regarding health attitudes and behavior. Nevertheless, research evidence indicates that most mass media campaigns oriented toward changing health care habits fail. The objectives of this paper are to analyze why health care campaigns fail and to derive generalizations for more effective use of mass media by health care professionals.

Overview

The role of mass media in affecting knowledge, attitudes, and behavior toward health care may be thought of in terms of the following two dimensions.

1. Mass media may impact health knowledge, attitudes and behavior both in a deliberate sense through "campaigns" that are specifically designed for such impact, and in an unintended or "incidental learning" sense through material that contains health-related information, but which is not specifically intended to impact health knowledge, attitudes or behavior.
2. In both cases, mass media may act either as a "change agent" or as a "reinforcing agent" -- that is, media may function in such a way as to change knowledge, attitudes and behavior or to confirm existing behavior patterns. In these respects, the role of mass media in affecting health care is similar to their role in affecting knowledge, attitudes and behavior toward other products and services.

Campaign Versus Unintended Effects

Mass media campaigns are intended to communicate certain health care information with a view toward change in health habits. Examples include anti-smoking, seat belt usage, lower cholesterol, and hypertension identification campaigns.

Mass media may also have unintended effects in the sense that the average viewer is exposed to a regular diet of "medical" shows on television and also to large numbers of commercials for proprietary medicines. The learning from such programming and commercials may be in the form of "misinformation" and may not be compatible with good health habits. A national study by the Louis Harris Organization (1973), for example, concluded that mass media were second only to the individual's physician as a source of health information. Furthermore, much of the health information absorbed from television is likely to be under low involvement conditions and, therefore, processed without evaluation.

A logical question then is whether mass media depict an accurate profile of health, illness, and the value of medical services, drug products, or medical treatment. Some social critics suggest that mass media depict a distorted and stereotyped view of these topics with consequences for people's health beliefs,

attitudes and behavior and for their probabilities of accessing the medical system under specified conditions. For example, to what extent does advertising for proprietary drugs convince people to search for simplistic solutions to medical symptoms that may be indicative of more serious problems? To what extent does cigarette advertising help people to deny or sublimate the medically dangerous effects of smoking?

The extent to which mass media either positively or negatively impact health is an important empirical question requiring systematic evidence to resolve. One study of television programming found that 30% of the health-related information was "useful" while the remaining 70% was inaccurate or misleading or both (Smith, 1972). This may suggest the magnitude of the potential problem, although this study is only one isolated piece of research evidence. Another study by Frazier et al. (1974) of dental health advertisements concluded that 43% of the information is inaccurate, misleading, or fallacious. The hypothesis may well be that mass media act more to misinform than to educate people about health and appropriate health habits.

Change Agent or Reinforcing Agent

The potential of mass media communications in the health care arena is generally phrased in terms of their promise for changing habits and life styles. However, the history of communication research indicates that the most persistent finding is that mass media act mainly to reinforce existing attitudes and behavior.

The ability of mass media to effect change is actually a function of a number of factors and requires certain conditions which we will develop later in this paper. Basically, however, the probability of change tends to be a function of how much commitment people have to existing behavior patterns. Under high commitment conditions, as is frequently the case in health care, bringing about change may indeed be a difficult undertaking. This is likely to be the case since health behavior is frequently rooted both in long term reinforcement patterns and in support by the individual's social environment. (In some special cases physical and psychological addiction patterns may also be a factor with which to contend.) A look at the evidence on health care campaigns supports the statement that most health care campaigns do not succeed among large numbers of intended subjects. The literature is replete with discouraging case studies.

- Obesity. In summarizing the evidence on obesity, Stunkard (1975) sets forth five propositions: (1) most obese people do not enter treatment, (2) of those who do, most drop-out, (3) of those who remain, most do not lose much weight, (4) of those who lose weight most will regain it, and (5) many of those entering treatment pay a high emotional price. Nevertheless, Stunkard registers considerable hope based on behavior modification programs, which recently have improved the treatment of obesity. He implicitly rejects mass media as an important force in changing behavior.
- Smoking. Anti-smoking campaigns have had limited success, at best. Cigarette consumption has not

declined, despite communication campaigns and public policy initiatives protecting non-smokers. In fact, it is increasing among teenagers, especially among girls. However, there has been a change toward consumption of lower tar and nicotine cigarettes. Perhaps the consequence of messages about lower tar and nicotine cigarettes has been to convince smokers that smoking is becoming safer.

On the other hand, one potentially successful anti-smoking campaign was initiated, when counter-advertising messages were shown on television under the equal time provision of the Federal Communications Commission. Possibly the combination of smoking and counter-smoking commercials presented together acted similarly to a two-sided communication; however, it is unlikely that mass media counter-advertising alone accomplished the job.

Seat Belts. In a review of research on seat belt usage campaigns, Leon Robertson et al. (1974) report a general lack of positive results. These authors then initiated a well-controlled experimental study using split-cable television whereby one audience received messages advocating seat belt use and a matched audience on the other half of the cable did not receive messages. After a nine month period tracking actual seat belt usage behavior, the authors could only conclude that: "The campaign had no measured effect whatsoever on safety belt use" (p. 1077).

Community Fluoridation Programs. Despite endorsement by the United States Public Health Service and the Surgeon General, controlled fluoridation of community water supplies has more often been rejected than accepted by voters. Between 1950 and 1969, 1139 communities voted on fluoridation; the issue lost in 666 communities and won in 473 (HEW, 1970). One part of the difficulty is the complexity of the fluoridation issue and another part is voters' susceptibility to the fear appeals used by opponents.

Health Maintenance Organizations. Despite the advantages claimed for the HMO concept, enrollment campaigns have met with limited success--with a few notable exceptions (primarily the Kaiser plan). Perhaps the HMO concept is not as desirable as its advocates claim (Glasgow, 1972) or perhaps the benefits to consumers are not readily apparent and communication campaigns have underestimated the difficulties of changing medical behavior patterns.

Heart Disease. The most encouraging results on a mass media campaign are from the Stanford study conducting a program to reduce susceptibility to heart disease among residents of three communities. Instructional programs used in conjunction with mass media have documented attitudinal and behavioral changes in diet and cigarette smoking. The role of mass media alone in one community on a delayed continuity basis is almost as effective as the personal instruction-mass media combination (Maccoby and Farquhar, 1975). The cost-effectiveness of this campaign, however, is very much in question.

campaigns fail. These reasons may be summarized as follows:

1. Most health care campaigns operate without explicit objectives or with inappropriate or unrealistic objectives, probably because they are based on an inadequate understanding of the way mass communications work, and on an inadequate understanding of the marketing requirements of the "product" being promoted.
2. Most health care campaigns are non-programmatic; they are short-run, one-time efforts, while the behavior change they are designed to induce must continue in the long run.
3. The beneficial effects of the recommended behavior change are not immediately apparent to the consumer, and perhaps never will be.
4. Most health care campaigns fail to identify market segments within the total audience who require different communication approaches in line with their specific needs.

Setting Objectives and Assigning a Role to the Mass Media

It is not sufficient to seek knowledge change or attitude change without a mechanism for also achieving behavior change and it is difficult for the mass media to achieve behavior change. For example: most smokers have knowledge of the ill effects of smoking and many have a negative attitude toward smoking. Therefore, presenting them with more knowledge as to the negative effects of smoking is unlikely to have much impact. Instead, a communication campaign must be linked to a behavior change mechanism other than mass communications (such as behavior modification group enrollment) if the campaign is to be successful. But behavior change even so induced is unlikely to persist in the long run unless its beneficial effects are continuously reinforced, since the beneficial results from the behavior change are not apparent in the short run and since there may also be some gratifications attached to the previous behavior.

Although non-smoking is a regularly repurchased "product," this is a different marketing situation from the usual consumer packaged goods situation in which advertising is used to achieve trial, and in which reinforcement from use of the product is a significant force in accomplishing continuing use of the product. An important function of mass communications in changing health behavior must be to reinforce new behavior, since use of the "product" is insufficient reinforcement in itself. Fortunately, this is a role which mass media have continually demonstrated an ability to perform well. Nevertheless, behavior change must be accomplished first, and by means other than mass media.

As we have noted, most health campaigns are short-run, start and stop efforts, with little long-term systematic and programmatic planning. Yet, changing health is likely to involve both multiple channels of persuasion and regular long term reinforcement. In summary, people must be moved through a decision sequence which is likely to take some significant amount of time and different means of persuasion may have complementary and cumulative impact, and may be necessary to achieve persistent behavior change.

Segmentation

Most health care campaigns try to reach everyone. Yet, not all segments of the market are as likely to change and different segments may require different incentives for change. It is incumbent on the change

Why Health Campaigns Fail, and How to Help Them Succeed
Analysis of the foregoing and other campaigns indicates that there are some basic reasons why most health care

agent to specify the market segments likely to be receptive to change and to expect that different messages focused on different needs may be effective with different demographic and psychographic segments.

Examination of needs by segment may be mandatory. For example: smoking provides gratification for smokers; it fulfills certain needs. These needs may relate to anxiety patterns or may be tied to social interaction patterns (Wortzel and Clarke, 1977). Programs to help reduce smoking, we might argue, should help find alternatives for the continued satisfaction of these needs. Basically, if we are to change health we must do so in line with people's needs. It does little good to scare people, insult people, etc., except under certain extreme conditions. Changed health patterns must be shown to be in line with the audience's self-perceived needs.

Segmentation is also critical if mass media are used to support a campaign in which behavior change has been accomplished by other means. Reaching the yet unchanged will be wasteful, if not counter-productive in light of possible future efforts. It is essential to reach the changed in order to reinforce the behavior changing mechanism.

Conclusion: Principles for Health Care Change

Following is a set of tentative propositions for the successful design and implementation of health care campaigns.

1. Mass media communication by itself may be effective in initiating change, but generally only if the change sought is minor and consumers have low information needs. This is seldom the case in health care.
2. Mass media communication will generally be most effective at an early point in the health decision change process whereas personal sources will generally be most effective later in the decision change process.
 - a. Mass media communication objectives, therefore, must be tied toward encouraging people to access the professional health care system, or to sensitize them to other sources of communication.
 - b. The peer and professional system will constitute the subsequent supporting mechanism necessary to bring about actual change.
3. It is the cumulative effect of a communication campaign that eventually results in behavior change.
 - a. This indicates the need for repetition and reinforcement over time. Reinforcement of health change is a particularly important role for mass media.
 - b. This indicates the need for multiple information sources which play complementary roles -- including advertising, personal selling, peer support, and professional intervention.
4. Peer sources (personal influence) will be a particularly important source of legitimation and "reality testing" when the benefits of change are not obvious or cannot be demonstrated in the short run. This is likely to be the case in much of health care.
5. Communication campaigns for health care -- even when they make use of donated public media time

and space -- are not free. The real cost is the opportunity cost if the communication campaign could have been more successful.

6. A health care communication campaign must explicitly recognize the problem of selective perception -- that those who see the message may be those who are already concerned about the issue and engaging in recommended change activities.
7. A communication campaign may have to provide support for change within a family or peer group context. Obesity, for example, may be tied to family diet habits and change may depend on family involvement.
8. Communication messages must be keyed to the needs of the market segment being reached. It is necessary to offer positive alternatives and not simply to denigrate the individual's existing health habits.
9. Low returns should be expected in a communication campaign. Most advertising and persuasion seeks small levels of change -- in the range of 3 to 5 percent of the audience per year. Mass conversion in the short-run is indeed a rare phenomenon.

References

- Frazier, P. Jean, Joanna Jenny, Ron Ostman and Charles Frenick, "Quality of Information in Mass Media: A Barrier to the Dental Health Education of the Public," Journal of Public Health Dentistry, Vol 34, Fall, 1974, pp 244-257.
- Glasgow, John M. "Prepaid Group Practice as a National Health Policy: Problems and Perspectives," Inquiry, Vol. 9, September 1972, pp 3-15.
- Harris, Louis Organization, A study cited in The Report of the President's Committee on Health Education, New York, 1973.
- HEW, Fluoridation Census 1969 (Bethesda, Md: U.S. Dept of HEW, 1970).
- Maccoby, Nathan and John W. Farquhar, "Bringing the California Health Report Up to Date," Journal of Communication, Vol 26, Winter, 1976, pp 56-67.
- Maccoby, Nathan and John W. Farquhar, "Communication for Health: Unselling Heart Disease," Journal of Communication, Vol 25, Summer, 1975, pp 114-126.
- Robertson, Leon S., Albert B. Kelley, Brian O'Neill, Charles W. Wixom, Richard S. Eiswirth and William Haddon, "A Controlled Study of the Effect of Television Messages on Safety Belt Use," American Journal of Public Health, Vol 64, November, 1974, pp 1071-1080.
- Smith, Frank A., Geoffrey Trivaz, David A. Zuehlke, Paul Lowinger and Thieu L. Nghiem, "Health Information During a Week of Television," New England Journal of Medicine, Vol 286, March 2, 1972, pp 516-520.
- Stunkard, Albert J., "Presidential Address-1974: From Explanation to Action in Psychosomatic Medicine: The Case of Obesity," Psychosomatic Medicine, Vol 37 May-June 1975, pp 195-236.
- Wortzel, Lawrence H. and Roberta Clarke. "Environmental Protection for the Non-Smoker: Consumer Behavior Aspects of Encouraging Non-Smoking," Advances in Consumer Research, Vol V, Association for Consumer Research, 1978.

A REVIEW OF FORMAL THEORIES OF CONSUMER SOCIALIZATION

Deborah L. Roedder (student), Northwestern University
Nicholas M. Didow (student), Northwestern University
Bobby J. Calder, Northwestern University

Abstract

Consumer socialization research presently overrelies on Piaget's work as a source of theory. There is a need to bring other formal theories to bear. However, in using other theories researchers must be aware that socialization theories vary along a number of dimensions. This paper provides a framework for comparing formal theories. Three classes of theories are described in terms of this framework.

Introduction

As an area for empirical investigation, consumer socialization has largely developed in response to public policy issues surrounding children. Under the rubric of consumer socialization, research has mostly focused on age-related differences in children's attention perception, and evaluation of television commercials (e.g. Wartella and Etna, 1974; Ward, 1972; Ward, Levinson and Wackman, 1972; Blatt, Spencer, and Ward, 1972). Some attention has also been directed toward the interaction between parents and children in purchasing situations and the influence of the family in contributing to the development of the child's consumer skills (e.g. Ward and Wackman; Ward, Wackman and Wartella, 1977).

Since the focal point of research has been similarly public policy questions, the theoretical development of consumer socialization has proceeded in a rather ad hoc manner. Most investigators have adopted Piaget's theory of cognitive development as an explanatory framework. There has been little questioning of the suitability of Piagetian theory to the explanation of advertising effects, let alone its usefulness as the sole approach to investigating the wide range of phenomena which should be included under consumer socialization.

While there has been occasional criticism of the reliance on Piagetian theory (e.g. Calder, Robertson, and Rossiter, 1976), there has been almost no work exploring alternative theoretical sources. To date only Robertson and Feldman (1976) have addressed this critical issue.

As a starting point toward a more complete theoretical approach, Robertson and Feldman (1976) suggest that the area of consumer socialization should be divided into component problem areas which can then be matched with various theoretical approaches. They argue that available theories, such as learning theory and psychoanalytic theory, should be analyzed according to their problem orientation and utilized for the particular research problem under investigation.

As a blueprint for theoretical development, however, Robertson and Feldman's proposal offers a rather limited, unidimensional analytical framework. Although the proposal documents the overreliance on Piagetian notions and correctly identifies the need for incorporating alternative theories into consumer socialization research, the basis for analyzing these alternative approaches is limited to problem orientation. However, since the classical approaches to socialization were developed and extended by researchers in different disciplines, the resulting theories differ on a multi-dimensional basis. The theoretical formulations vary

not only in problem orientation, but in their basic orientation regarding socialization as a phenomenon.

The following discussion provides a multi-dimensional framework for analyzing alternative formal theories. The analytical framework developed is then used to compare these classes of formal theory. The objective is to go beyond problem orientation to specify more precisely the relevance of formal theory to consumer socialization research.

A Multi-Dimensional Framework for Analyzing Formal Theories of Socialization

The selection of a particular formal theory directs, and thus limits, the questions which a researcher can pose and seek to answer. This selection should be made with full appreciation of the dimensions upon which formal theories of socialization vary.

These dimensions are identified and organized here in a framework consisting of the following major categories: (1) the "process" orientation; (2) the "content" orientation; and (3) the "goal" orientation. These three categories and their specific dimensions are summarized in Figure 1 and explained in the discussion which follows.

FIGURE 1
Analytical Framework for Interpretation
of Formal Theories of Socialization

PROCESS ORIENTATION

passive view of the individual	_____	active view of the individual
external control over socialization	_____	internal control over socialization

CONTENT ORIENTATION

socialization viewed as cumulative	_____	socialization viewed as transitional
socialization emphasized through childhood	_____	socialization emphasized through life
emphasis on individual difference	_____	emphasis on group commonalities

GOAL ORIENTATION

descriptive focus	_____	explanatory focus
complete approach to socialization	_____	fragmentary approach to socialization
decision processes	_____	all processes

Process Orientation. Socialization theories differ in how they depict the basic process of socialization. After a review of the socialization literature, Zigler and Child (1969) suggest that one of the most fundamental

differences with regard to current theory is the view of the individual as a passive or active participant in the socialization process. One can view socialization as a process whereby social forces are applied to an individual or as a collaboration between society and the individual. A related issue concerns the control over the progression of socialization. That is, whether external conditions exert control over socialization or whether internal processes largely determine the progress made in socialization.

Content Orientation. Although the majority of socialization literature focuses on children and adolescents, some attempt has been made to study socialization at later stages in the life cycle (cf. Brim, 1966). In terms of our analysis, formal theories vary in their ability to encompass socialization processes at different stages in the temporal life cycle.

Another distinction can be drawn between perspectives that approach socialization as either cumulative or transitional in nature. Cumulative approaches assume that socialization follows some definitive path, that is, what has been learned in the past is a pre-condition or contributes heavily to future socialization. In contrast, transitional approaches view the socialization process as a less progressive movement from a beginning to an identifiable end.

A final content difference in formal theories is their emphasis upon individual differences versus group homogeneity in the socialization process. Although this dichotomy is most apparent in psychological approaches as opposed to sociological/anthropological approaches, some distinction is still possible in the viewpoints analyzed here.

Goal Orientation. The most visible difference between theories in terms of goal orientation is the focus of the conceptualization in terms of an explanatory versus descriptive approach. The descriptive approach tends to identify changes in the individual as a result of socialization, whereas the explanatory approach concerns itself with the actual process of change.

A less definitive distinction can be drawn between theories that offer a complete approach to socialization as opposed to those that offer only a fragmentary perspective. In this analysis, a complete approach is described as one capable of explaining the socialization process across a wide variety of situations. Fragmentary approaches are theories which provide a description or explanation of the process in limited contexts.

The final dimension is decision orientation. That is, whether or not the process of socialization involves the conceptualization of the individual as a decision maker.

The dimensions within each of the three major categories of the framework are more fundamental than any orientation toward specific kinds of problems. We will attempt to show that these dimensions have important implications for consumer socialization research. Three important classes of formal theory are analyzed in turn using the dimensions. These classes of theories are selected by the authors because, in the case of cognitive development theory, prominence in extant investigations of consumer socialization or, with regard to the latter two classes, because of their promise for future research. Other, more marginal theories (e.g. social anthropology, psychoanalysis, psychoanalytically oriented social anthropology, normative-maturational, genetic and constitutional) are described in Zigler and Child (1969).

The orthodox position of cognitive development theory is best represented by the work of Jean Piaget (cf. Piaget, 1950; Piaget, 1968; Piaget and Inhelder, 1969). Piaget's theory of cognitive development is based in his paradigm of the intellectually mature human. To Piaget the fully developed person represents a paradigm of universal order and equilibrium between the interactive parts and the individual as a whole (Maier). Thus, development is a function of the drive for equilibrium and equilibrium depends on activity and experience. The active view of the person continuously interacting with the environment in search of equilibrium is the basic Piagetian model. This interaction, which is controlled primarily by the individual, identifies previously unknown contradictions and absences in the cognitive structures of the person (Zigler and Child, 1969). The person seeks a balance with the environment, yet such a balance can only be obtained by a new equilibrium through the processes of assimilation and accommodation.

In this view, the environment of the neonate is initially undifferentiated. How the social and ideational environment becomes differentiated for an individual in cognitive development is a result of the socialization of the person (Maier, 1969) or the manifestation of phylogenetic drives as experienced during the formative cognitive years of early life. The person seeks to adopt, or to find equilibrium in the environment, by assimilation of events into the cognitive structure, while simultaneously accommodating the cognitive structure to the requirements of the events (Calder *et al.*, 1976). The interpretation of the events as one would conceive them (assimilation) and the modification of the cognitive structure by the actual event (accommodation) are complementary cognitive processes. The interaction of these two processes over the stages of cognitive development is the source of the summary description of Piaget's theory of the person as "ever changing, ever the same."

The content of cognitive development theory hypothesizes a cumulative view of changing cognitive structure and behaviors, as opposed to any transition content. Piaget's concern for cognitive development spans the chronological range from birth to cognitive maturity, which usually occurs in late adolescence. The emphasis is upon the typical individual, rather than individual differences, groups, institutions, or cultures.

Cognitive development theory hypothesizes "unity in continuity" (Maier, 1969, p.4) in both quantitative and qualitative behaviors. Existing cognitive structures and behaviors are highly dependent on previous ones. The developmental schedule of Piaget represents a cumulative series of sequential phases in which each later phase is related to previous acquisitions. The qualitative differences in the person's cognitive processes are differentiated into Piaget's multiple stages (cf. Piaget and Inhelder, 1969). These stages of different cognitive processes do not vary in sequence. One's progress through the stages may be accelerated or retarded only somewhat by various environmental factors of socialization (Kohlberg, 1971). Another perspective of this cumulative process is that it represents a continuous unfolding of generalization and discrimination, with each cognitive structure representing to some degree a repetition of the structures of previous stages in a slightly different form (Maier, 1969). This unfolding continues until the point of developmental maturity. Cognitive developmental maturity is achieved in most cases in the mid-to-late adolescent years (Piaget and Inhelder, 1969).

Piaget focuses on the development of cognitive thought, or the intellectual aspect of personality development,

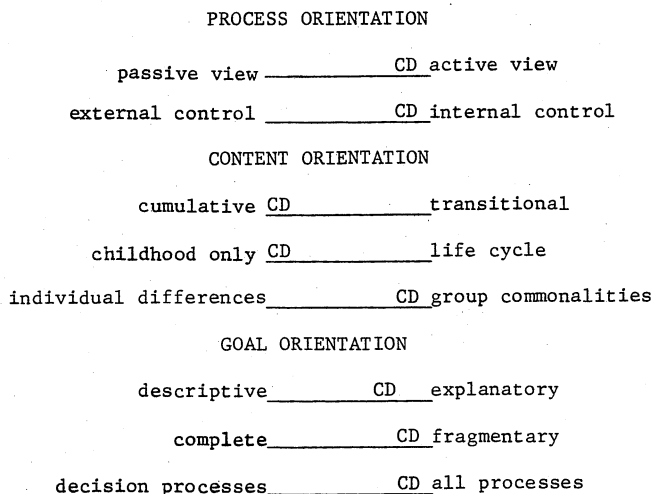
as opposed to affective or behavioral development. In this respect Piaget's theory of cognitive development is not intended to be a complete explanation of the concepts one would include in the nominalogical network of socialization.

The theory of Piaget is a description of marked developmental changes in cognitive processing over a fixed succession of stages (Zigler and Child, 1969). These changes include both qualitative reorganization of cognitive processes over the stages and the emergence of new cognitive structures in each stage. Developmental changes are different from maturational changes. The former reflect sociopsychological changes, while the latter represent only biological changes and organic growth. The successive stages of cognitive development are identifiable by developmental phases of readiness and acquisition, not by maturational phases or chronological time intervals (Maier, 1969). From the chronological perspective, however, the age at which a child achieves a particular stage is a function of previous cognitive ability and subsequent experiences (Reese and Lipsett, 1970).

The focus of Piaget's theory of cognitive development is not necessarily that of decision-making. Rather, Piaget is concerned with the individual's competencies and limitations with regard to comprehension of a situation. Cognitive development theory does not directly focus on normative choice situations which are one of the concerns of traditional consumer socialization research. On the contrary, the emphasis of Piaget is on those constructs of cognitive processes which are related to general, hypothetical levels of intellectual development.

The boundaries of cognitive development theory identified by the analytical framework in Figure 2 are cause for concern in many respects. The most severe limitation of Piaget's theory is its restricted emphasis upon childhood and early adolescence. By strict adherence to this perspective, we deny many areas of potential interest throughout an individual's life, such as consumer socialization issues connected with stages of family formation and dissipation, the development of family consumer "decision-making styles," and the "retirement community." If notable changes in consumer behavior occur within the individual's lifetime, a theory of consumer socialization should be capable of encompassing such phenomena.

FIGURE 2
Summary of Analysis of Cognitive
Development Theory (the symbol CD
positions the theory on each dimension)



A related problem concerns the viewpoint of consumer socialization as cumulative in nature. This of course stems from Piaget's perspective of cognitive development as occurring in stages towards the final goal of the individual's equilibrium with the physical world. For the purposes of an adequate formation of consumer socialization, it seems appropriate to include transitional learning in the theoretical perspective. Certainly we cannot ignore the identity transformations that individuals undergo during their lifetime, which have implications for changing consumer perspectives. An example is the transformation from a teenager to a bachelor to a husband to a retiree. Because Piaget is interested in the physical environment, it also becomes problematic to deal with the symbolic content marketing phenomena, such as product images.

The adaptation of Piaget's cognitive development theory in consumer socialization research focuses investigation upon the child's limitation in sensation and cognitive perceptual ability. Cognitive development theory addresses long-term sociopsychological changes. It is of marginal help in explaining the specifics of cognitive responses (Calder et al, 1976), short-term fluctuations (Robertson and Feldman, 1976), or the affective aspects of intelligence and behavior (Robertson and Feldman, 1976).

Finally, Piaget's theory does not address the question of discriminative interpretation of the environment and does not explicitly view the individual as a "decision-maker". The child is viewed as searching to understand the environment. Perceptual boundness is included in the analysis, yet the individual's selective processing of information is characterized as a cognitive deficiency rather than a decision process that is so important in an information processing perspective. Only if one recognizes the limitations of process, content, and goal, can Piaget's theory of cognitive development be of benefit in consumer socialization research.

Traditional Learning Theory

Almost all definitions of socialization share at least one common element. This general element is variously expressed in verb form as "acquires," "develops," or "transmits". What the element means is that socialization results in a relatively permanent change in the behavior of an individual. The generic term for the relatively permanent change is "learning". Where the unanimity of "learning" is lost, though, is with regard to the differences in process, content, and goal of the various formal theories of learning.

This section reviews the general characteristics of formal connectionist learning theories and formal cognitive learning theories. Each theory individually and in combination makes a substantial contribution to understanding the general area of socialization. Because of the breadth and all-inclusive nature of the construct of learning, some authors suggest that socialization is synonymous with a social learning process. This perspective combines psychology, sociology, and anthropology into the traditional research of socialization (Goslin, 1971).

Hill (1971) argues that the field of learning theory may be initially divided into connectionist theories and cognitive theories. Learning theories in the connectionist tradition are represented by the stimulus-response contiguity theories of Watson and Guthrie and by the reinforcement theories of Thorndike, Skinner, and Miller. The theories of Lewin and Tolman represent the traditional cognitive positions. More contemporary connectionist shifts toward the cognitive positions are represented in theories by Spence and Mowrer. Cognitive social learning theories are formal attempts at synthesis of the

two historic divisions.

Hill (1971) identifies the boundaries of the division of connectionist and cognitive when one theory is taken independently of the other to explain learning in humans. On the one hand, when dealing with humans it is probably myopic to refer to all learning in the connectionist mode of the increased or decreased probability that certain responses will occur. Human responses can only be understood in the context of beliefs, attitudes and goals. Likewise, the pure cognitive mode, such as Lewin's life space, omits any interactive mapping between the environment and the life space of the individual. It is the lack of holism of the first perspective and the inoperable nature of the second perspective which are addressed by a third group of synthesis theories such as cognitive social learning (cf. Mischel, 1973; Mischel, 1976; Bandura, 1971).

Process for connectionist learning theory is the behavior change which results from classical conditioning and operant conditioning. Learning by classical conditioning happens first. It occurs before the child has the capabilities of language and classification. After language development begins, learning by operant conditioning involving problem solving, decision-making, and other cognitive contingency skills, is feasible in addition to continued involuntary classical conditioning (Reese and Lipsett, 1970).

Connectionist learning theories differ from cognitive learning theories along the several process dimensions. In general, connectionist theories view the process as one which is externally controlled and which occurs as effectively if the person is active as it does if the person is passive. Cognitive learning theories conversely view the person as an active participant in the process. For them, internal cognitions and beliefs have a significant personal impact on the process.

Within the connectionist perspective classical conditioning stimulus-response theories address learning situations in which the outcomes are dependent on externally imposed reward contingency schedules. In this perspective an external agent has complete control over the process. Instrumental conditioning theories focus on learning which occurs due to rewards and punishments which result from active behavior of the person (Reese and Lipsett, 1970). In both instances the control process of reward or nonreward is external to the person involved in the learning situation.

The cognitive perspective hypothesizes cognitive representations and active manipulations involving contingencies which may be quite complicated. In the traditional cognitive learning theories the representations and contingencies are subject to influence from external sources, but the learning process is itself dominated by the person.

Generally, learning theory has no temporal limitations. Relatively enduring changes in behavior which may occur at any age are the focus of the content. The one modification to this statement reflects the cognitive developmental stage theories (cf. Kohlberg, 1971), which are incorporated within some of the more contemporary cognitive-developmental learning theories. These theories are thus focused more on learning in the pre-adolescent years, the years during which the cognitive processes of the person proceed through the fixed hierarchical sequence. Social learning theories are also broad enough to accommodate variations at different developmental levels (Zigler and Child, 1969).

Learning theories vary with respect to whether the content is cumulative or transitory. Connectionist learning theories may be imposed at any point in any phenomenon. The more cognitive learning theories, and especially those which also include developmental

aspects of cognition, are more cumulative in content. Cognitive social learning theories, Mischel's for example, require a cumulative idiographic content understanding in order to make any reliable predictions of behavior.

Connectionist learning theories focus exclusively on group commonalities of learning. Contiguity theorists and reinforcement theorists are concerned with the general conditions under which the response will become more or less probably. Early cognitive theorists likewise address commonalities of cognitive representations and contingency development. Contemporary cognitive social learning theory is idiographic rather than nomothetic. The researcher is thus able to investigate and seek to understand individual differences in learning which are contained within the theoretical perspective of the phenomenon. Bandura and Walters (cf. Zigler and Child, 1969), for example, recognize that different cognitive contingencies and available social models within a single age stratum result in considerable interindividual variations.

The final area in which characteristics of learning theory are considered is that of the goal of the theory. Formal learning theories are attempts at complete theories of explanation which are not limited solely to decision-making phenomena.

In the perspective of learning theory all behavior, purposeful or not, is synonymous with learning. In this view, all socialization is synonymous with social learning in a general sense (Goslin, 1971). The difficulty with this idea is that each of the two major areas of traditional learning theory, connectionist and cognitive, are found to in fact be incomplete theories when taken independently. The synthesis positions of cognitive social learning theories are thus intuitively appealing. The controversy, though, centers upon the lack of existing assessment techniques to provide consistent, meaningful measures of the theoretical constructs when they are applied to substantive areas like consumer socialization. In this sense even the synthesis positions currently fall short of their goal.

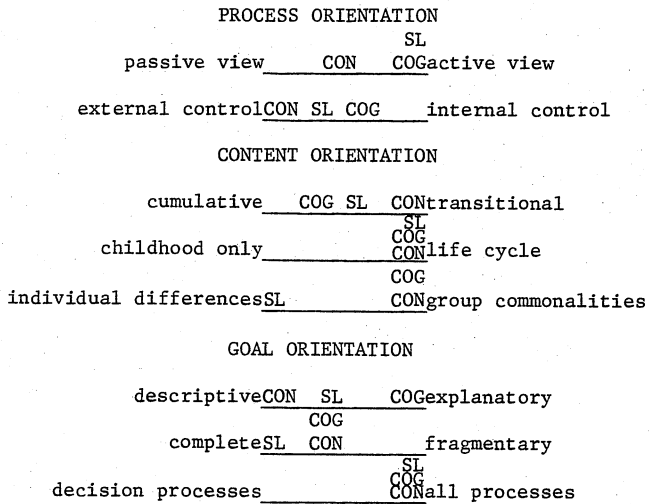
Mischel (1973) argues that a complete formal theory of behavior processes must be inclusive of aspects of the person, the situation, and the interaction of the person and the situation. His cognitive social learning theory involving person variables of cognitive competencies, encoding strategies and perceptual constructs, symbolic exchange expectancies, differential personal values, and self-regulatory internal and external goals and plans is an encompassing attempt at a complete theory of explanation and prediction. Operationalizing the constructs of Mischel in substantive socialization areas presents quite a challenge, however.

Learning theories, even with their inherent boundaries as identified in the analytical framework of Figure 3 should be very useful in many substantive areas of socialization research (see Zigler and Child, 1969, 465-468). Hopefully, our discussion renders the range of these theories clearer.

Interaction Theory

A great deal of theorizing and empirical work in sociology and social psychology could be subsumed under the general heading of "interaction" theory. For the purposes of this analysis, the discussion of interaction theory will center around symbolic interaction and role theory. It is indeed difficult to make distinctions between these two perspectives, since many role theorists reflect a symbolic interactionist viewpoint in some form or another. However, most role theory departs somewhat from symbolic interaction. It reflects

FIGURE 3
Summary of Analysis of Learning Theories
(the symbols CON, COG, SL position
connectionist, cognitive, and social
learning theories respectively)



a partially deterministic viewpoint different from symbolic interaction.

Symbolic interaction theory attempts to explain how an individual acts or thinks in relation to an environment which is symbolic as well as physical in nature. Man is characterized as actively seeking and learning meaning, values, and behaviors through interaction with others. The individual learns about himself through encounters with other people and learns to "predict" the behavior of others by taking the role of another individual (individual other) or taking the role of a number of persons acting together in some group or society (generalized others).

Role theory has adopted the notion of Mead's role-taking in the symbolic interaction process as its point of departure and major focus of elaboration. The general proposition of most role theorists is that overt behavioral and attitudinal phenomena can be codified into meaningful categories called roles. From this starting point, role theorists have developed quite divergent conceptualizations - the idea of role has been used to denote prescription to norms and demands, enactment and involvement in role performance, overt and covert processes, and behavior which the individual initiates versus that which is directed toward the person (Biddle and Thomas, 1966). The more social psychologically oriented theorists have described social interaction by the mutual sharing of role expectations whereas many sociologically oriented theorists have tended to use role concepts in describing an individual's integration into society.

On the basis of this brief description, it should be clear that most interaction theory stresses an active view of the individual in social processes. Symbolic interaction theorists have been consistent in their view that the individual constructs the "world as taken for granted", the symbolic environment, through interaction with other people. Role theorists have been somewhat less consistent in viewing the individual as an active participant. Some observers (e.g. Linton, 1947) present the view that roles are externally constructed without any individual input. Other writers also see roles as socially structured, yet stress the importance of the individual's participation in learning, performing, and discriminating roles in the process

of socialization. An extreme position in terms of the individual's participation is presented by Turner:

"The idea of role-taking shifts emphasis away from the simple process of enacting a prescribed role to devising a performance." (Turner, 1962, p.23.

In terms of external control, symbolic interactionists recognize that the individual's environment is an integral influence in the process of learning and interaction. However, they strongly resist the notion that individuals or things in the environment control the behavior or the socializing individual (Blumer, 1962).

In contrast, some of the proponents of role theory see the process as a limited form of social determinism - focusing more on social influences. The behavior of the individual is formed by the sanctions, rules, and demands of others in addition to input from the individual regarding his own understanding and prescriptions of what his behavior should be (Biddle and Thomas, 1966). Other writers, most notably Sarbin and Allen (1968), totally reject any suggestion of a conformity theory and reflect the symbolic interactionist emphasis upon individual input and activity.

In terms of generality, both symbolic interaction and role theory are capable of studying socialization throughout an individual's life. Both viewpoints recognize that the individual changes group affiliations and confronts new environments in the course of a lifetime. With regard to the symbolic environment, man is constantly creating and changing meanings:

"Socialization continues throughout life in another sense: the society and its groups are constantly creating new meanings and values: (Rose, 1974, p. 147).

Role theorists are even more explicit in their concern for socialization issues throughout the life cycle. Some roles are characterized as being nonrepetitive in nature - in effect, they are "stages" in the life cycle. Other roles that are successively enacted, such as wife and mother, must also be learned at different stages in the life cycle.

Both interactionist positions also share the view that socialization is more transitional in nature than cumulative. Although general skills are needed for effective role enactment and interaction (such as cognitive and motor skills), the specific skills are usually learned at the time the individual enters a new role situation (the obvious exception being anticipatory learning). As Sarbin and Allen (1968) note, most of the general skills are learned in childhood and contribute to later role learning. But, the specific skills learned for a particular role are usually not transferable to another position - especially roles that are nonrepetitive in nature. Successive roles to be learned do not usually build upon each other and may often conflict.

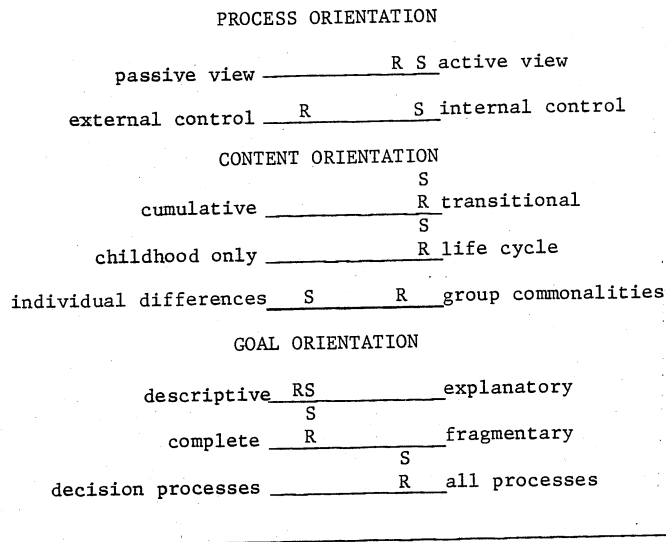
Regarding aggregative emphasis, role theorists tend to look more towards group (in this case role) characteristics than do symbolic interactionists. Roles are usually perceived to be structured by some measure of group consensus; however, role expectations do differ between individuals in the same position. For symbolic interactionists, the interaction process proceeds through a "definition of the situation" agreed upon by the participants; for role theorists, this is accomplished through the sharing of role expectations.

Although these interaction approaches appear to offer rich conceptualizations for socialization processes, they fail to adequately explain the learning processes

involved in socialization. Symbolic interactionists have been more interested in dynamic and developmental processes than role theorists who focus largely on description and structures. The sparse amount of work which deals with role learning suggests that individuals develop conceptions of role enactment through social learning processes, both incidental learning and explicit coaching (Strauss, 1966).

Interaction theories thus offer an active view of the individual. They treat socialization throughout the life cycle, and within many contexts (see Figure 4).

FIGURE 4
Summary of Analysis of Interaction Theories
(the symbols R and SI position role and symbolic interaction theories respectively).



Conclusion

We have developed a framework for the interpretation of formal socialization theories and have positioned three classes of theories within the framework. Analysis of the theories reveals marked variation along the dimensions of the framework. Selection of formal theories for consumer socialization research must thus be viewed as more complex than a unidimensional choice dictated by a problem orientation. Rather, theory selection requires careful consideration of the multidimensional nature of theories. Although a theory may seem applicable on one dimension, it may also dictate the researcher's view of socialization on other dimensions.

The diversity of theories of socialization indicated by our analysis also points up the weakness of an off-the-shelf approach to explanation in consumer socialization research. Ultimately, the goal must be to contribute to the development of theory. This can best be done by comparing the explanatory power of existing theories.

References

Albert Bandura, "Social Learning Theory of Identification Processes," in D.A. Goslin (Ed.), Handbook of Socialization Theory and Research. Chicago: Rand McNally and Co., 1971

Bruce Biddle and Edwin Thomas (Eds.) Role Theory: Concepts and Research. (New York: John Wiley and Sons, Inc., 1966).

J. Blatt, Lyle Spencer, and Scott Ward, "A Cognitive Developmental Study of Children's Reaction to Television Advertising," in E.A. Rubinstein, G.A. Comstock, John P. Murray (Eds.), Television and Social Behavior. Volume III. Television in Day to Day Life: Patterns of Use. (Washington, D.C.: Department of Health, Education, and Welfare, 1972, 452-468).

Herbert Blumer, "Society as Symbolic Interaction," in A.M. Rose (Ed.) Human Behavior and Social Processes. (Boston: Houghton Mifflin Co., 1962).

O.G. Brim, Jr., "Socialization Through the Life Cycle," in O.G. Brim and S. Wheeler, Socialization After Childhood. New York: John Wiley and Sons, Inc., 1966).

Bobby J. Calder, Thomas Robertson, and John Rossiter, "Cognitive Response to Advertising: The Relation of Child to Adult Models," in Beverlee B. Anderson (Ed.), Advances in Consumer Research, Volume III. (Cincinnati: Association for Consumer Research, 1976).

D.A. Goslin, Handbook of Socialization Theory and Research. (Chicago: Rand McNally and Co., 1971).

Winifred F. Hill, Learning - A Survey of Psychological Interpretations. (Scranton: Chandler Publishing Co., 1971).

Lawrence Kohlberg, "Stage and Sequence: The Cognitive-Developmental Approach to Socialization," in D.A. Goslin (Ed.) Handbook of Socialization Theory and Research. (Chicago: Rand McNally and Co., 1971).

Ralph Linton, The Study of Man. New York: Appleton-Century, 1936, 114. See also Ralph Linton, "Concepts of Role and Status," in Newcomb and Hartley (Eds.) Readings in Social Psychology. (New York: Henry Holt, 1947).

Henry W. Meier, Three Theories of Child Development. (New York: Harper and Row, Publishers, 1969).

Walter Mischel, "Toward a Cognitive Social Learning Reconceptualization of Personality," Psychological Review 80 (1973): 252-283.

Walter Mischel, Introduction to Personality. (New York: Holt, Rinehart, and Winston, 1976).

Paul H. Mussen, John J. Conger, and Jerome Kagan, Basic and Contemporary Issues in Developmental Psychology. (New York: Harper and Row, 1975).

Jean Piaget, The Psychology of Intelligence. (New York: Harcourt and Brace, 1950).

Jean Piaget, Six Psychological Studies (New York: Random House, 1968).

Jean Piaget, and Barbel Inhelder, The Psychology of the Child. (New York: Basic Books, Inc., 1969).

Hayne W. Reese and Lewis P. Lipsett, Experimental Child Psychology. (New York: Academic Press, 1970).

Thomas S. Robertson and Shel Feldman, "Children as Consumers: The Need for Multitheoretical Perspectives," in Beverlee B. Anderson (Ed.), Advances in Consumer Research, Volume III. (Cincinnati: Association for Consumer Research, 1976).

Thomas S. Robertson and John Rossiter, "Children and Commercial Persuasion: An Attribution Theory Analysis," Journal of Consumer Research. (June, 1974: 13-20).

Arnold M. Rose, "A Summary of Symbolic Interaction Theory," in R. Serge Denisoff, Orel Callahan, and Mark H. Levine (Eds.) Theories and Paradigms in Contemporary Sociology. (Itaska, Illinois: F.E. Peacock Publishers, Inc., 1974).

Theodore R. Sarbin and Vernon L. Allen, "Role Theory" in G. Lindzey and E. Aronson (Eds.) Handbook of Social Psychology, Volume II. (Cambridge, Mass.: Addison-Wesley, 1968).

Anselm Strauss, "Coaching," in Bruce J. Biddle and Edwin J. Thomas (Eds.) Role Theory: Concepts and Research. (New York: John Wiley and Sons, Inc., 1966, 50-54).

Ralph H. Turner, "Role-Taking: Process Versus Conformity" in Arnold M. Rose (Ed.) Human Behavior and Social Processes. (Boston: Houghton Mifflin, Co., 1962).

Daniel B. Wackman and Scott Ward, "The Development of Consumer Information-Processing Skills: Contributions from Cognitive Development Theory," in Beverlee B. Anderson (Ed.), Advances in Consumer Research, Volume III. (Cincinnati: Association for Consumer Research, 1976).

Scott Ward, "Children's Purchase Influence Attempts and Parental Yielding," Journal of Marketing Research 9 (August, 1972): 316-319).

Scott Ward, "Children's Reactions to Commercials," Journal of Advertising Research 12 (April, 1972: 37-45).

Scott Ward, "Consumer Socialization," Journal of Consumer Research 1 (September, 1974: 1-14).

Scott Ward, David Levinson, and Daniel Wackman, "Children's Attention to Television Advertising," in E.A. Rubinstein, G.A. Comstock, John P. Murray (Eds.), Television and Social Behavior, Volume III. Television in Day to Day Life: Patterns of Use. (Washington, D.C.: Department of Health, Education, and Welfare, 1972: 491-516).

Scott Ward, Daniel Wackman, and Ellen Wartella, "The Development of Consumer Information-Processing Skills: Integrating Cognitive Development and Family Interaction Theories," in William D. Perreault, Jr., (Ed.), Advances in Consumer Research, Volume IV. (Atlanta: Association for Consumer Research, 1977: 166-171).

Ellen Wartella and James S. Ettma, "A Cognitive Developmental Study of Children's Attention to Television Commercials," Communication Research, 1 (January, 1974, 44-69).

E. Zigler and I.L. Child, "Socialization," in G. Lindzey and E. Aronson (Eds.), The Handbook of Social Psychology (Second Edition), Volume 3, The Individual in a Social Context. (Reading, Mass.: Addison-Wesley Publishing Co., 1969).

CHILDREN'S CONSUMER INFORMATION PROCESSING: REPRESENTATION OF INFORMATION
FROM TELEVISION ADVERTISEMENTS

Ellen Wartella, The Ohio State University
Daniel B. Wackman, University of Minnesota
Scott Ward, Harvard University and Marketing Science Institute

Abstract

This paper considers several conceptual and methodological issues involved in examining how children represent TV information in their memory system. It is suggested that researchers examine children's constructed memory for advertising information as well as their verbatim memory for commercial elements. The kinds of interpretations and connections children make regarding the nature of the advertised product and its relationship to other similar products are proposed as ways in which children might "construct" meaning for the advertised information. Furthermore, the paper discusses how to measure children's memory for advertisements and suggests that both recall and recognition measures should be used. This is to insure that researchers do not underestimate young children's processing abilities.

Introduction

Models of adult consumer information processing have tended to distinguish those processes involved in viewer reactions to television advertising from those mental processes involved in product choice situations (Hughes and Ray, 1974). The activities of selecting, interpreting and making sense of advertising information are referred to here as representational processes since they concern how advertising information is represented in the child's cognitive system. This paper will focus on several conceptual and methodological issues surrounding the way investigators model children's representation of information from television advertising stimuli.

The notions presented here are an extension of previous work adopting a cognitive developmental perspective on children's consumer information processing by Ward, Wackman and Wartella (1977). The basic model of information processing outlined in this previous research has relied heavily on Piaget's theory of cognitive development to suggest major dimensions along which children younger and older than middle childhood vary. Perhaps best-known of these dimensions is that of perceptual boundedness: younger children's tendency to focus on the perceptual and surface characteristics of products and advertisements (Wartella and Ettema, 1974; Ward and Wackman, 1973). We have viewed these dimensions of cognitive growth and the general cognitive abilities available to children as "rules for processing information" at various levels of cognitive development (Kohlberg, 1969).

This initial conceptual model followed rather strictly Piaget's theory and was useful in isolating the general distinctions among grade school children's knowledge about TV advertising and products. For instance, between kindergarten and sixth grade, children acquire greater understanding of the purpose of advertising; they tend to select both more and varied kinds of information when recalling a television commercial and describing a product, and gradually use multiple attributes when comparing brands of a product group (Ward, Wackman and Wartella, 1977; chapters 3 and 4).

In our current research, our interest is to go beyond these general dimensions of children's consumer information processing and elaborate more fully various as-

pects of how children represent information from specific TV advertisements. Our research is still strongly cognitive developmental in perspective but less closely aligned to strict Piagetian formulation.

Watching Television as an Information Processing Task

In our previous research we examined children's information processing transituationally; i.e., the same mental processes intervening between the input of a stimulus and the child's output of a response were assumed to occur both in the television viewing situation and in product choice situations. Consequently, we tended to focus on information processing activities which were similar in both task situations. For example, information selection from an advertisement was operationally defined as recall of elements in a commercial; and information selection about products was operationally defined as the kind of attributes children focus upon when they are asked to consider buying a new product (Ward, Wackman and Wartella, 1977). For purposes of describing general characteristics of children's thought about the consumer environment, this conceptual and methodological approach proved useful, as suggested above. However, in order to increase conceptual clarity of those activities involved in consumer information processing, it seems advisable to examine these two task situations independently. In this way, more specificity in points of divergence and similarity between the two processing tasks can be examined.

Furthermore, as Dawes (1975) points out, any model of information processing is necessarily circumscribed by the task being modeled. In particular, problem-solving tasks might best be viewed as conceptually distinct from non-problem solving tasks (Berlyne, 1970). While many of the same mental activities are involved in both situations, the tasks posed for the child information processor are probably much different.

For instance, a major distinction between the television viewing situation and the product choice situation is the degree to which each task is problem-oriented, where in the child is seeking information to reach some solution or decision, such as a product choice. It would seem reasonable to assume that in most instances when young, grade-school children sit down in front of a TV set and watch a commercial there is little "planfulness" involved in how they select information and little "intention" to seek information to use in a purchase decision. This assumption is based on the evidence of relatively low comprehension of the purpose of advertising by grade school children and only moderate awareness of TV advertising as a source of new product information across a wide range of products (Ward, Wackman and Wartella, 1977).¹ Such directed and planful use of

¹ For instance, among kindergarten, third and sixth grade subjects surveyed, only four percent, 15 percent and 38 percent respectively, indicated that they recognized the selling motive when they responded to the question: Why are commercials shown on television (Ward, Wackman and Wartella, 1977; Table 4-4, p. 60). Further, when these same children were asked to name sources of new product information for toys, clothes and snack

television advertising probably occurs only rarely, and most likely at particular times during the year, such as Christmas time when children are seeking gift ideas (Caron and Ward, 1975). On the other hand, when children are asked to "choose" a cereal at the store or are given money to spend on a product of their choice, the information processing task at hand involves directed thinking activities and intentional use of information to solve a problem, i.e., to "buy" the product that best satisfies the child's needs or desires.

This is not to say that children don't use information which they have learned from television commercials to reach a purchase decision; indeed, there is evidence that television viewing influences product requests at Christmas time (Robertson and Rossiter, 1977). This suggests that viewing TV advertising may directly affect product choice by influencing the child to buy a particular product, i.e., by raising the salience of product X above all other brands of a product group. Alternatively, TV advertising may influence the child's strategy for approaching a product decision, by suggesting certain attributes of brands in a product class to be considered, i.e., suggesting that the child buy the brand of a product class that has the most of attribute X (Wright and Barbour, 1975).

However, the crucial point is that television advertising most likely enters product decision-making tasks sometime after viewing the television advertisement even if during TV commercial viewing the child's desire for a product is raised. This suggests that we should examine children's memory of what they have seen and heard from television commercials since it is what they have stored and retrieved from memory which will influence their processing activities at the time of product decision-making.

We are interested here, then, in exploring how the information presented on the television screen is stored and retrieved from the child's memory system. Rather than focus on models of information processing which are directly problem-solving task-oriented, such as Newell and Simon (1972) and Pascual-Leone (1969, 1970), we have chosen to examine models of memory development for further conceptualization of how children select, interpret and comprehend TV advertising information.

Representation of Information in Memory

The view of memory adopted here distinguishes between two types of memory: episodic memory, or memory for a specific event which occurred at a specific time and place; and semantic memory, or the accumulated knowledge one has acquired about the world (Brown, 1975). Piaget and Inhelder (1973) refer to the former as "memory in the strict sense" and the latter as "memory in the wider sense." Episodic memory involves memory for directly experienced occurrences, the actual input or verbatim recollection of experience and "for discrete perceptual instances that are distinct and separable from the larger unit in which they occur" (Brown, 1975; p. 136). Thus, episodic memory is usually what is referred to as verbatim memory for a television commercial or program, as it involves remembering the elements of the television commercial. On the other hand, Brown (1975) notes that semantic memory involves "memory for meaningful systems of units in context." Such memory is constructive and holistic, and it is memory for the gist of a narrative or story, such as the overall "message" a viewer constructs from a TV commercial.

foods, fewer than one-third of the kindergarteners, and only about one-half of the third and sixth graders mentioned TV commercials (Tables 4-2; p. 57).

Further, as Brown (1975) points out, no particular interaction with an environmental stimulus is totally one type of memory experience or the other for a child; aspects of both episodic and semantic memory are involved whenever a child interacts with the environment. Children's semantic memory system, in the broadest sense their acquired knowledge about the world and their attendant cognitive abilities, skills and language, helps determine what they will "remember" about any specific episode or occurrence, i.e., it influences children's episodic memory:

"What the head knows has enormous effect on what the head learns and remembers. . . Older individuals will presumably store, retain and retrieve a great many inputs better or differently than younger ones, for example, simply because developmental advances in the content and structure of their semantic or conceptual systems render these inputs more familiar, meaningful, subject to gap filling or otherwise more memorable for them. (Flavell, 1977; p. 189)

This perspective maintains that memory is a constructive process. Memory "involves an imaginative reconstruction or construction built on extant knowledge" (Brown, 1975). Both at the point of storing and retrieving information from memory, the subject is constructing and reconstructing an internal representation of that information to be remembered (Paris, 1975).

One result of this constructive aspect of memory is that children attempt to integrate information they remember to comprehend the "gist" of the stimuli presented. Paris (1975) reviews several studies in which children in second and fifth grade were presented with a series of sentences which told a story. Later the subjects were given a series of sentences to read, some of which they had actually read earlier, some of which were new but preserved the meaning of what they had read, and some of which were new sentences that did not preserve the meaning. Paris reports that the children consistently confused the original sentences with those new sentences which preserved the correct meaning of the narrative. This finding (which has also been found for memory of pictorial stimuli) suggests that children integrate semantic information to construct a holistic meaning of that information. They go beyond the information given to integrate ideas and form inferences. Further, as this research points out, children's integration of information is sometimes at the expense of correct recognition of information they had actually seen or heard.

The perspective on memory outlined above has a very basic implication for research on children's information processing of television advertisements: researchers should examine more than verbatim or episodic memory for particular elements of commercials, such as brand name or product attributes mentioned. Examination must be made of the kinds of inferences and connections children make when integrating the advertised information into their semantic memory system, i.e., what overall message do the children take away from the commercial? As Paris' (1975) work suggests, even in the absence of children's faithful and accurate retrieval of specific elements from an advertised message, these children may still "remember" some constructed or integrated meaning from the message. It may be this "constructed meaning" from the message which children recollect from the advertisement during product decision-making situations.

In our current research, we are exploring several ways in which children might go beyond the information given in the TV commercial, including: (1) by drawing connections between the product and themselves, such as how they can use the product or what will happen to them after buying the product; (2) by making comparisons

between the product advertised and other brands of that product class, such as how game X is the same as or different from other games the child has played; (3) by making inferences about the people and activities shown in the commercial to arrive at inferred attributes of the product. Researcher attention to such connections and inference-making should further elucidate cognitive characteristics of children's representation of TV advertising information. This is particularly the case since past research on information processing of advertisements has primarily examined children's selection and recall of commercial elements (Ward, Wackman and Wartella, 1977).

A second implication of memory research for consumer information processing studies concerns the kind of developmental differences in processing activities which may be observed. Verbatim memory of commercial elements and constructed memory for the commercial's meaning may not show the same type of developmental effects. Brown (1975a, 1975b) argues that where memory tasks involve primarily episodic memory, the researcher should expect a "levels difference" on the measure of memory retrieval. For instance, various research studies have indicated that as children grow older they recall a greater number of elements (Ward, Wackman and Wartella, 1977); thus the level of recall performance increases with age. However, where memory retrieval tasks engage primarily the semantic memory system, Brown predicts that the researcher will find a pattern difference, some interaction of developmental level with the task variables. For instance, one might predict a "patterns difference" in the type of inferences children in grade school make about the television commercial message. Considering that children younger than middle childhood have been shown to have difficulty making comparative judgments about objects (Cellerier, 1972), such children might be less likely than older children to make connections between the product advertised and other brands of that product class. Older children might be more likely to make multiple kinds of connections and inferences about the advertising information.

This discussion of the kinds of constructed meanings children might make of advertising information and the types of developmental effects which might be hypothesized for children's representation of advertising information should serve to illustrate the relevance of memory research for conceptualizing children's consumer information processing. Furthermore, the literature on memory development has implications for the methodologies employed to measure processing activities.

Measures of Representational Processes

Measurement of children's memory for television advertising information has typically employed open-ended recall measures (Rossiter, 1975; Rubin, 1972; Ward, Wackman and Wartella, 1977). Furthermore, researchers have focused primarily on verbal memory to the neglect of visual memory (Rossiter, 1975). Although the particular modality used to code information in memory is still open to debate (cf. Bransford and McCarrell, 1974; Brown, 1975; Piaget and Inhelder, 1973), it seems reasonable to allow for the possibility of multiple forms of representation--visual, verbal and imaginal. As Rossiter (1975) has suggested, multiple memory codes should be measured.

The issue of how best to measure retrieval of information from memory may be a thornier issue than previous research has acknowledged. Particularly, more attention should be paid to the distinctions between recall and recognition memory. While researchers may tend to think of recall and recognition questions as measurement tools, recall and recognition activities also constitute types

of retrieval activities for the subject being interviewed. Moreover, recall and recognition memory place different task demands on the subject for actively retrieving information from memory.

Recognition involves external memory cues for the child, such that there is already something present in immediate experience to assist the retrieval process. There are no such external cues present in recall memory. In recall, the subject has to do more of the retrieval job himself: "recall is the more difficult process as it demands regeneration in the absence of the stimulus" (Brown, 1975, p. 111). Piaget (1968) and Brown (1975) have proposed that there is a developmental progression in the development of these retrieval activities, such that recognition memory develops earlier in the child than does recall memory. Several studies provide evidence that recognition is a more efficient memorial process than recall for children younger than middle childhood (A.L. Brown, 1975b; Ritter, et al, 1973; Kobasigawa, 1974). This research literature suggests that both recognition and recall measures should be used to examine young children's memory for television advertising information. It may be the case that previous research has underestimated what children are learning from television advertising. In our current research we are exploring this possibility.

A series of experiments are in preparation to examine kindergarten and third grade children's information processing of several specially produced commercials for hypothetical brands of candy and game products. A post-viewing questionnaire is under construction to measure children's recall and recognition of both discrete visual and verbal elements of the commercials and various kinds of inferences and connections children might make to integrate the commercial information. The general question strategy is to begin each type of question with a recall measure (e.g., "What product did the commercial show you?") and then to follow that open-ended recall question with a series of multiple-choice recognition items (e.g., "Was it: a) a raisin candy b) a chocolate covered raisin c) a chocolate covered peanut?"). Through a revolving sequence of recall and visual and verbal recognition measures we are attempting to tap children's episodic and semantic memory for the product advertised and the TV commercial narrative message.

Pretest data collected in Columbus, Ohio on 18 kindergarten and 12 third grade children's episodic memory for advertised information about a chocolate covered raisin candy are presented on the next page. Recall and recognition measures were made after one TV commercial exposure. Children's memory for the type of product advertised, brand name, and two attributes mentioned about the candy was measured, first by a recall measure and then if the answer was either incorrect or incomplete, by a recognition measure. As the table indicates, third graders show better memory overall of the advertised product than do kindergarteners, particularly when recall measures are employed. However, when recognition measures are used, the number of kindergarteners "remembering" the commercial elements improves considerably. Although these data refer to episodic memory, we are also utilizing recall and recognition measures to examine

2 Piaget (1968) introduces an intervening retrieval process called reconstruction memory which occurs when the child is presented with the elements of a pattern he was previously shown and is then asked to reconstruct the correct pattern from the elements. Piaget suggests that the developmental sequence of memory development moves from recognition to reconstruction to recall memory. Reconstruction memory seems less useful as a general measurement tool for advertising research and thus has been ignored here.

children's semantic integration of the advertised information.

Table 1

Number of Children Recalling or Recognizing Information About a Chocolate Covered Raisin Candy After One Exposure to a TV Advertisement for that Product

	<u>Kindergarten</u>	<u>Third Grade</u>
<u>Memory for Product Type</u>		
Recall Measure	6	12
Recognition Measure	$\frac{8}{N = 18}$	$\frac{-}{N = 12}$
<u>Memory for Brand Name</u>		
Recall Measure	3	10
Recognition Measure	$\frac{9}{N = 18}$	$\frac{2}{N = 12}$
<u>Memory for Primary Product Attribute</u> (raisins have vitamins)		
Recall Measure	3	7
Recognition Measure	$\frac{7}{N = 18}$	$\frac{1}{N = 12}$
<u>Memory for Secondary Attribute</u> (# of candies in box)		
Recall Measure	-	5
Recognition Measure	$\frac{5}{N = 18}$	$\frac{-}{N = 12}$

The suggestion that recognition measures provide better estimates of children's verbatim memory for commercial messages is a rather obvious statement. Surprisingly, such recognition items are rarely employed. If the goal of the researcher is to identify optimally how much of the commercial message children do store in memory, then recognition measures should be used.

One argument which might be advanced for not using recognition measures rests on an implicit assumption regarding how advertised information is used in children's product decision-making. If it is assumed that information which can be recalled, i.e., self-generated without external cues, has the greatest impact on subsequent decision-making, then recognition memory of a TV advertisement would seem relatively unimportant. Alternatively, if it is assumed that children's product decision-making at least sometimes occurs in situations - where children do have cues to remind them about the advertised information, e.g., product choices which occur in stores, then recognition memory for advertised information may be relatively more important in product decision-making. However, until we know more about how

and where children do reach decisions about product choice, and the kinds of situational factors which are present during decision-making, it is difficult to assess the validity of these alternative assumptions. Our current research strategy is to measure both recall and recognition memory since we are interested in assessing how much of the advertised information children do store regardless of which retrieval strategy is used in subsequent decision-making.

To summarize our position, we believe that research on children's information processing of television advertisements can be informed both conceptually and methodologically by considering the development of memory activities. Greater attention should be paid to semantic memory or how children integrate advertising messages into their general knowledge system. By examining how children construct meanings of the advertised information and represent that meaning in their memory system we hope to acquire greater understanding of how children use advertising information in subsequent product decision-making. This research perspective should lead to greater specificity in conceptualizing and measuring children's information processing activities.

References

D.E. Berlyne (1970) Children's reasoning and thinking. In P.H. Mussen (ed.) Carmichael's Manual of Child Psychology, third ed. New York: Wiley.

J.D. Bransford and N.S. McCarrell (1974) A sketch of a cognitive approach to comprehension: some thoughts about what it means to comprehend. In W.B. Weimer and D.S. Palermo (eds.) Cognition and Symbolic Processes New York: Winston.

A.L. Brown (1975a) The development of memory. In H.W. Reese (ed.) Advances in Child Development and Behavior, 10 New York: Academic Press.

A.L. Brown (1975b) Recognition, reconstruction and recall of narrative sequences by preoperational children. Child Development, 46, 156-166.

A. Caron and S. Ward (1975) Gift decisions by kids and parents. Journal of Advertising Research, 15, 4, 15-20.

G. Cellier (1972) Information processing activities in recent experiments in cognitive learning--Empirical studies. In S. Farnham-Diggory (ed.) Information Processing in Children New York: Academic Press.

R.M. Dawes (1975) The mind, the model and the task. In F. Restle, R.M. Shiffrin; N.J. Castellan, H.R. Lindman and D.P. Pisoni (eds.) Cognitive Theory, Vol. 1. New York: John Wiley and Sons.

J.H. Flavell (1971) First discussant's comment: What is memory development the development of? Human Development, 14, 272-278.

J.H. Flavell (1977) Cognitive Development. Englewood Cliffs, N.J.: Princeton-Hall.

G.D. Hughes and M.L. Ray (1974) Buyer-Consumer Information Processing. Chapen Hill, N.C.: University of North Carolina Press.

A. Kobasigawa (1974) Utilization of retrieval cues by children in recall. Child Development, 45, 127-134.

L. Kohlberg (1969) The cognitive developmental approach to socialization. In D. Goslin (ed.) Handbook of

Socialization Theory and Research. Chicago: Rand McNally.

- D.E. Liebert, J.N. Sprafkin, R.M. Liebert and E.A. Rubenstein (1977) Effects of television commercial disclaimers on the product expectations of children Journal of Communication, 27, 118-124.
- A. Newell and H.A. Simon (1972) Integration and inference in children's comprehension and memory. In F. Restle, et al, (eds.) Cognitive Theory, Vol. 1. New York: Wiley, 223-246.
- J. Pascual-Leone (1970) A mathematical model for the transition rule in Piaget's developmental stages. Acta Psychologica, 63, 301-345.
- J. Pascual-Leone and J. Smith (1969) The encoding and decoding of symbols by children: A new experimental paradigm and a neoPiagetian model. Journal of Experimental Child Psychology, 8, 328-355.
- J. Piaget (1968) On the Development of Identity and Memory. Worcester, Mass.: Clark University Press and Barre.
- J. Piaget and B. Inhelder (1973) Memory and Intelligence New York: Basic Books
- K. Ritter, B.H. Kaprove, J.P. Fitch and J.H. Flavell (1973) The development of retrieval strategies in young children. Cognitive Psychology, 5, 310-321.
- T.S. Robertson and J.R. Rossiter (1977) Children's responsiveness to commercials. Journal of Communication, 27, 101-106.
- J.R. Rossiter (1975) Visual and verbal memory in children's product information utilization. In B.B. Anderson (ed.) Advances in Consumer Research, Vol. 3. Association for Consumer Research, 523-527.
- R.S. Rubin (1972) An exploratory investigation of children's responses to commercial content of television advertising in relation to their stages of cognitive development. Unpublished Ph.D dissertation, University of Massachusetts.
- D.B. Wackman and S. Ward (1975) The development of consumer information processing skills: contributions of cognitive development theory. Advances in Consumer Research, Vol. 3. Association for Consumer Research.
- S. Ward and D.B. Wackman (1973) Children's information processing of television advertising. In P. Clarke (ed.) New Models for Communication Research. Beverly Hills, California: Sage Publications. 119-146.
- S. Ward, D.B. Wackman and E. Wartella (1977) How Children Learn to Buy: The Development of Consumer Information Processing Skills. Beverly Hills, Ca.: Sage Publishing.
- E. Wartella and J.S. Ettema (1974) A cognitive developmental study of children's attention to television commercials. Communication Research, 1, 46-69.
- P. Wright and F. Barbour (1975) The relevance of decision process models in structuring persuasive messages. Communication Research, 2, 246-259.

THE EFFECTS OF TV MESSAGES FOR HIGH AND LOW
NUTRITIONAL FOODS ON CHILDREN'S SNACK AND
BREAKFAST FOOD CHOICES

Marvin E. Goldberg, McGill University
Gerald J. Gorn, McGill University
Wendy Gibson, (student) Stanford University

Abstract

This study investigated the capacity for food messages developed for television to influence the nature of children's snack food and cereal selections. In a controlled experiment, children were exposed to varying numbers of commercials for highly sugared snack foods and cereals, or Public Service Announcements (PSA's) for less sugared, nutritious fruits and vegetables. In addition, a 24 minute edition of the 'Fat Albert' program starring Bill Cosby called 'Junk Food' was utilized. After exposure to one or another of these stimuli, children were presented with a number of snack food and cereal choices. Exposure to the Fat Albert program (even when it was interrupted by the commercials for snack foods) was most effective in reducing the number of highly sugared foods selected and increasing the number of fruits and vegetables selected. Exposure to the PSA's or the snack food commercials was also capable of creating differences in the foods selected by the children. These differences were noted even though children in all the conditions were aware of which foods were "healthy" or "unhealthy".

In April 1977, the consumer activist group, Action for Children's Television (ACT) petitioned the FTC to prohibit the advertising of candy to children on television. At just about the same, the office of Research and Analysis of the New York State Assembly made virtually an identical recommendation to the New York State Legislature related to the advertising of all heavily sugared products to children (Mauro and Feins, 1977).

The focus of the discontent with present advertising practices revolves around the postulate that given their relatively underdeveloped cognitive abilities, young children are likely to be unduly influenced by the more than 20,000 TV commercials that they are exposed to each year (Adler et al, 1977); the majority (55%) of which are for either breakfast cereals or snack foods (Gussow, 1972). ACR (1977) has cited evidence from associations such as the American Dental Association and the Society for Nutrition Education to document the fact that frequent sugar consumption in the form of the typical pre-sweetened breakfast cereals and snack foods usually advertised on TV contributes especially to tooth decay in children, as well as other potential physical problems such as obesity. While there seems to be relatively little controversy regarding the potentially negative effects of an overconsumption of sugar and sugared products, what appears to remain an issue of greater contention is the establishment of a causal link between exposure to children's TV advertising for sugared snacks and breakfast foods and subsequent attitudes and behavior regarding these food products.

Correlational evidence (Clancy-Hepburn, Hickey and Nevill, 1974) is suggestive of such a link, as is observational data (Atkin, 1975; Galst and White, 1976), and experimental evidence (Goldberg and Gorn, 1977; Poulos, 1975). These notwithstanding, a recent summary of the research in this field concludes that the nature and extent of these effects is not clear and that much re-

mains yet to be assessed (Adler et al, 1977).

While the majority of both policy-makers and researchers have tended to focus on the degree to which TV commercials affect children adversely, consideration has more recently also been given to the manner in which TV can be used to encourage positive or pro-social behavior, including sound dietary habits in children. Thus a variety of groups have been involved in developing Public Service Announcements and program material that encourage children to eat nutritious foods such as fruits and vegetables and to moderate their intake of highly sweetened snacks and cereals they see advertised on TV (see Society for Nutrition Education, 1975). The White House Conference on Food, Nutrition and Health (1970) specifically recommended that the President establish a permanent task force to formulate a mass-media nutrition education program. More recently the New York State Office of Research and Analysis recommended in their report (March, 1977) that stations be requested to air PSA's emphasizing proper nutrition for children, to be paid for by a tax levied on all broadcasters as a function of their advertising revenues. These efforts suggest that at least these policy makers and other interested parties believe there is a link between the material children are exposed to on TV (messages featuring either low or high nutritious foods) and their subsequent attitudes and behavior.

Research has documented that: very young children in particular are most likely to believe television commercials (Ward and Wackman, 1973); are most likely not to understand disclaimers (Atkin, 1975), and are least likely to understand the sales intent of the commercials (Rubin, 1972; Robertson & Rossiter, 1974). Thus television may have an especially strong impact on the very young, with the potential to maximally influence their food habits which for these children may not be as yet fully formed.

One of the most elaborate nutrition education efforts via the mass media is a series of programs called "Mulligan Stew". It includes a series of six half hour films integrated with written teaching materials that attempts to get children to learn the various nutrients associated with the four basic food categories. An experiment evaluating the effectiveness of this program revealed that it was successful in communicating the nutritional information but was, on the whole, unable to get the children to change their consumption behavior (Jenkins et al, 1975).

The present study sought to test the potential of TV as a medium for influencing children's snack and breakfast food preferences. The 30 second message - either Public Service Announcement for highly nutritious foods or commercials for less nutritious foods was one vehicle used. The second was a 24 minute program entitled "Junk Food", one of a number of issues treated by a series called "Fat Albert" starring Bill Cosby who notes:

What this series attempts to do is harness the natural interest (enjoyment of cartoons) and act as a conveyor of learning to instill an awareness of life and develop codes of behavior.

Employing animation as the primary medium to teach its (Fat Albert's) goals differs from Sesame Street and the Electric Company in that its emphasis is upon the affective development of young viewers (1976, p. 60-61).

It was anticipated that exposure to these materials, insofar as they made the various foods attractive and salient, would serve as cues to the children and encourage the selection of either more or less nutritious foods, as the case might be. It was also considered likely that an entire program devoted to nutrition would be more effective than the 30 second PSA's. The increased time would not only direct attention to nutritional issues for a longer period of time, but potentially ought to allow for greater creativity in developing a meaningful theme.

In a further effort to vary the degree of exposure to messages for more or less nutritious snack and breakfast foods, conditions were developed in which the number of PSA's or commercials either approximated the number permitted per half hour by the NAB Code or was double that number.

Method

Subjects

The twin objectives of utilizing children who were as young as possible, yet at the same time could respond to simple paper and pencil measures resulted in the selection of a sample of first grade children.

Subjects were drawn from three Northern California-Bay area schools, about one mile from one another. The communities surrounding the schools were part of the same upper-middle class suburban area and were essentially indistinguishable from one another. In all, there were seven first grade classes in the three schools (three in one school and two in each of the remaining two schools). Total enrollment in those classes was approximately 175 of whom 122 or 70% participated in the study, (having returned signed permission statements from their parents). Subjects were randomly allocated to one of eight conditions (described below). At each school, assistants would draw a randomly selected subset of children from different classes and take them to a room where the experiment was conducted. Typically seven or eight would then watch one of the following professionally videotaped and edited programs.

Independent Variables

1. "Yogi's Gang", a 24 minute animated cartoon program provided the context for a 30 second commercial for each of the nine food products listed below. These were typically inserted, two at a time at five appropriate points in the program, for the total of 4.5 minutes of commercial time; an amount roughly corresponding to what children typically see in a half hour on childrens TV (N.A.B. Code, 1975). Commercials were for: Mounds Candy Bar, Lollipop Lifesavers, Crackerjacks, Hershey Candy Bar, Blow Pops, Milky Way Candy Bar, Kool-Aid, Sugar Crisps, and Fruity Pebbles.
2. "Yogi's Gang" plus the commercials for the nine food products, each repeated once for a total of 9 minutes of commercial time. These were inserted three at a time at six points during the program.
3. "Yogi's Gang" plus eight Public Service Announcements (for a total time of 4.5 minutes) stressing the values of eating fruits, vegetables, milk; and suggesting moderation in the use of candy bars,

sugared cereal and other sugared snacks. (The majority of these were selected from among those produced under the direction of the Bay area Committee on Children's Television; two produced by ABC were taped off-air). One or two of these Public Service Announcements were inserted at the same five points in the program as in condition 1.

4. "Yogi's Gang" plus eight Public Service Announcements each repeated once (for a total time of 9 minutes). These were inserted two or three at a time at the same 6 breaks as in condition 2.
5. A 24 minute "Fat Albert" program entitled "Junk Food". This is an animated childrens cartoon aired weekly on CBS, starring Bill Cosby. The program depicts "Fat Albert" and his friend "Slim" who eat only Junk Food (mostly non descript candy). This leads to a toothache for Slim followed by an anxiety producing visit to the dentist. This convinces Fat Albert, but not Slim, to turn to good healthy foods. Finally, Slim having skipped meals in favor of junk food is so weakened that he is responsible for losing the "big" football game. Most of the program is animated, with Bill Cosby appearing occasionally to underline a point made in the course of the animated cartoon. The theme of the program generally is: "This is Bill Cosby coming at you with music and fun and if you're not careful you may learn something before it's done."
6. The Fat Albert "Junk Food" program with the eight Public Service Announcements described above - each repeated once for a total of 9 minutes. In essence this condition represented the maximum message with regard to eating highly nutritious foods. The PSA's were typically inserted three at a time at six points in the program; (because of limitations of time and subject population we were unable to run this condition with the eight Public Service Announcements shown once only; i.e. for a total of 4.5 minutes).
7. The Fat Albert "Junk Food" program with the nine 30 second food commercials described above for a total of 4.5 commercial minutes. These were typically inserted two at a time at five points in the program. (Because this program was originally aired on CBS during Saturday morning, this condition comes close to replicating reality; i.e. what children viewing the program off-air would actually have viewed: the "Junk Food" program with its message to moderate candy and sugared product intake, ironically along with commercials for several examples of these sugared products.
8. A control group not exposed to any program at all. (Although the focus of this study was on the direct comparisons permitted by the groups described above, it was decided to incorporate this group to establish "baseline" data, as well as to assess the comparability of children in each of the three schools used in this study).

Dependent Variables

Following the program the children were involved in a four-five minute discussion of their favorite TV programs so as to break any momentary set that the program material might have established. At that point the children turned to the questionnaire that was used to obtain the dependent measures.

Subjects were first asked to indicate if they were a boy or girl and were then shown the first of a series of 3' x 2' boards each divided by lines into six of the six rectangles so that a child saw six actual snack foods

mounted on a board. Following consultation with nutritionists, three of the six foods were selected because they represented "healthy" highly nutritious snacks and the remaining foods were selected because they represented snacks that would be "unhealthy" and low in nutrition if over-eaten at the expense of other more nutritious foods. The first board had the following three snacks considered as highly nutritious: banana, peanuts, raisins; and three sugared snacks considered as less nutritious: Mounds Candy Bar, Jelly Beans and Lollipop Life-savers. The first page of the questionnaire had a pictorial representation of the food sketched on the page in the same location as on the board and identified in writing as well. The experimenter reviewed the items on the board in front of the children and subsequently on the first page of the questionnaire and made sure they recognized the correspondence between the board and the questionnaire page. No problems were encountered in this regard throughout the experiment.

The experimenter then structured the situation for the children's selection of the snack foods by indicating the following:

Now I want you all to pretend something. Let's pretend that your Mommy and Daddy were going away on a holiday, and they asked me to babysit for you while they were gone. Now I don't know the kind of foods you would want while they were gone. So suppose I said here are six snacks - you can eat three of them. Now you can tell me which three you would want by putting a big X through the three snacks on your page you would want most.

After they had indicated their three choices the experimenter removed the first board and revealed a second board with a second set of six snack foods; once again with three foods considered highly nutritious (apple, carrots, orange) and three considered less nutritious (Crackerjacks, chocolate "stars" and Ju-Jubes). The experimenter introduced the second board by suggesting:

Now let's suppose it was the second day I was babysitting for you while your Mommy and Daddy were away on a holiday. And again, because I didn't know the kind of snacks you like, I said here are six snacks and you can pick three of them to eat today.

After the children had once again made their three choices, the experimenter went on to the choices for the "third day of babysitting" (pear, cheese and crackers, fruit cocktail-snack cup; Hershey Candy Bar, Malted Milk Balls and Blow Pops) and then the "fourth day of babysitting" (strawberries, peaches-snack cup, raisins; Milky Way Candy Bar, Hershey Chocolate Kisses, Gum Drops).

It is important to note that in each set of six foods one or two of the sugared candy products were among the ones advertised in the commercial conditions, while one or two were more akin to the non-descript sugar products used in the Fat Albert program. Among the 3 more nutritious snacks presented in each set, one or two were among those featured in the Public Service Announcements while one or two were not.

Following the four boards of snack foods the experimenter went on to introduce two boards consisting of breakfast foods:

O.K. now what about breakfast. I wouldn't know what you would want for breakfast so you would have to tell me. Let's see, you could have:

Once again the six choices presented on the board consisted of three foods selected with the help of nutritionists: milk, orange, eggs and toast; and three considered less nutritious: Kool-aid, Super Sugar Crisps and Fruity Pebbles (sugared cereal). These foods were pictorially represented in the same fashion in their questionnaire and here too, the children were asked to "put an X through the 3 foods you would want for breakfast". Once completed, the children were provided with a second set of breakfast choices consisting of the same more nutritious alternatives (milk, orange, eggs and toast); and Cap'n Crunch, Frankenberries and Kool Aid as the less nutritious alternatives. (Only the cereals were changed in an effort to provide the same reasonably realistic breakfast choices in each set). While the four cereals are all roughly equally popular, the first two (Sugar Crisps, Fruity Pebbles) were among the nine advertised products in the commercial conditions while the latter two (Cap'n Crunch, Frankenberry) were not.

The second set of questions focused on the children's level of awareness of the healthy or unhealthy nature of each of the foods they just had seen on each of the six boards (four sets of six snack foods plus two sets of six breakfast foods). After all the children's choices had been made, the experimenter returned to the first board of snack foods and asked the children to turn to the next page of their questionnaire which once again included a pictorial representation of the foods corresponding to the first board. In this instance, however, each food had a "smiling face" with the word "good" and a "frowning face" with the word "bad" immediately below it. The experimenter indicated:

Now I want to ask you a different kind of question. Let's think about Do you think are good for you and healthy or bad for you and not healthy? If you think are good for you and healthy put a big X through the happy face that says good for you. If you think are bad for you and not healthy then put a big X through the sad face that says bad for you.

Now what about? Are good for you and healthy or bad for you and unhealthy? Put an X through the happy face that says good for you and healthy or an X through the sad face that says bad for you and unhealthy.

The experimenter continued in this way for each of the foods on each of the remaining boards.

At the conclusion of the entire set of questions, where time permitted, the experimenter engaged in a group discussion with the children, asking questions such as: why they had made choices they did; what they thought could be done given the "we all sometimes eat things that are not healthy for us."

Results

The children were able to select a maximum of either three more nutritious or three less nutritious foods on each of six boards, making a maximum of 18 possible in either direction. The control group selected a mean number of 10.2 less nutritious foods (and 7.8 more nutritious foods). As indicated in Table 1 these figures are very similar for the children in each of the three schools suggesting comparability on this key dimension.

TABLE 1

Mean Number of Less Nutritious Foods Selected by Children at Each School

School	N	\bar{x}	S
A	10	9.7	4.9
B	5	10.8	4.8
C	5	11.0	4.4

Initial examination of the mean scores revealed no significant difference between the groups exposed to either 4.5 minutes of PSA's ($\bar{x}=8.67$) or 9 minutes of PSA's ($\bar{x}=8.73$) in the context of the Yogi's Gang program. The difference between those exposed to either 4.5 minutes of commercials ($\bar{x}=12.12$) or 9 minutes ($\bar{x}=13.14$) was almost as small and equally insignificant. Given the similarity of data in the two groups exposed to Public Service Announcements and the similarity of data in the two groups exposed to commercials, it appeared more parsimonious to combine each pair in further analysis. [In fact, the similarity of the results with repetitive exposure parallels several earlier studies suggesting that repetition does not further change short term attitudes much beyond the impact of the initial exposure (Goldberg and Gorn, 1974, 1977)].

An analysis of variance conducted with the four original conditions and the two merged conditions described above, revealed highly significant overall differences ($F=9.62$, 5, 116 df, $p < .001$). Table 2 presents the means for each of the six groups in ascending order.

TABLE 2

Mean Number of Less Nutritious Foods Selected

Independent Variables	\bar{X}	S.D.	N
Fat Albert	2.87	4.03	15
Fat Albert & P.S.A.'s	5.60	5.78	10
Fat Albert & Com'ls	6.06	4.04	16
Yogi's Gang & P.S.A.'s	8.70	6.71	30
Control	10.70	4.55	20
Yogi's Gang & Com'ls	12.58	4.04	31

Post Hoc Newman Keuls analyses suggested the following significant differences between particular groups. Children exposed to Fat Albert selected significantly fewer less nutritious foods ($\bar{x}=2.87$) than those exposed to Yogi's Gang and PSA's ($\bar{x}=8.70$), who in turn selected significantly fewer than those who were exposed to Yogi's Gang and commercials ($\bar{x}=12.58$). The control group fell between the latter two ($\bar{x}=10.20$) but was not significantly different from either.

As might be expected, children who viewed Fat Albert with commercial insertions selected a somewhat greater amount of the less nutritious foods ($\bar{x}=6.06$) than those who viewed the program uninterrupted ($\bar{x}=2.87$) but this difference was not significant. Similarly, but perhaps more surprisingly, adding 9 minutes of PSA's to the Fat Albert program also tended to increase the number of less nutritious foods selected (5.60) but, again, not to a statistically significant extent.

In sum, while Fat Albert by itself was significantly more effective than Yogi's Gang plus PSA's in reducing the number of less nutritious foods selected, inserting commercials or Public Service Announcements into Fat Albert tended to reduce the effectiveness of the program to the level of Yogi's Gang plus PSA's. Nevertheless, children in each of these latter conditions still selected

ted significantly fewer less nutritious foods than the controls or those in the Yogi's Gang plus commercial condition.

An analysis of variance performed on just the 4 sets of snack food choices considered together was equally significant ($F=10.69$, 5, 116 df, $p < .001$). Post Hoc analysis revealed essentially the same pattern of results as were obtained for the total set of foods.

While the overall analysis of variance was also significant for the 2 sets of breakfast foods considered separately ($F=4.32$, 5, 116 df, $p < .001$), Post Hoc Newman Keuls revealed only one significant difference: the group seeing Fat Albert selected significantly fewer less nutritious breakfast foods ($\bar{x}=1.53$) relative to either the control group ($\bar{x}=3.45$) or the groups exposed to Yogi's Gang plus commercials ($\bar{x}=3.87$).

The second set of questions asked children to indicate whether each of the 36 foods they were presented with was "good for you and healthy" or "bad for you and not healthy". Very few errors were made in response to this question. In fact, the mean number of errors for all 122 subjects was 2.62. The overall analysis of variance revealed no significant differences between the groups ($F=1.12$, 5, 116 df, $p > .05$).

Discussion

The results of this study suggest that the type of snack and breakfast foods children select may be a function of the type of food messages they are exposed to on TV. Those who viewed material stressing the attractiveness and value of eating more nutritious foods (Fat Albert, PSA's) selected significantly more of these types of foods than those who viewed commercials stressing the attractiveness and value of eating less nutritious foods.

The Fat Albert program "Junk Food" was more effective than the PSA's in reducing the number of less nutritious foods selected. A number of explanations may be provided for this finding. For one, the unique constellation of creative factors associated with the Fat Albert program in general, and with Bill Cosby it's star in particular, is probably responsible in large measure for the program's effectiveness.

By contrast, at least some of the PSA's seemed less imaginative, slower moving etc. Of course, a program format with 24 minutes allows for the development of themes with antecedent events leading to subsequent consequences, while the 30 second PSA's make this approach difficult if not impossible. The question remains still, how effective could the Fat Albert approach be in a thirty second PSA?

The evidence in this study suggest that the degree of effectiveness of the various TV messages was not a matter of mere minutes of exposure. Doubling either the number of minutes of PSA's or the number of minutes of commercials had no increased effect on the children who were exposed to the messages. Further, adding 9 minutes of PSA's to the Fat Albert program did not enhance it's effectiveness. In fact, the number of less nutritious foods selected increased in this condition, even if not significantly so. Observation of the children in this condition suggested that the redundancy of the messages contained in the nine minutes of food PSA's added to the 24 minute "Junk Food" program resulted in the children paying little attention to the TV by the end of the sequence. This type of viewer satiation has been observed among adults who are overexposed to TV commercials (Greenberg and Suttoni, 1974) and among children as well (Goldberg and Gorn, 1974, 1977).

While Fat Albert appeared most effective in reducing the number of less nutritious foods selected, it should be emphasized that the Public Service Announcements were also able to significantly reduce this number (and concomitantly to increase the number of more nutritious foods selected). Of course real-life insertion of these PSA's would likely occur together with other commercials many of which might well be for candy, chocolate and other sugared foods. The impact of such "mixed" TV exposure remains to be tested. However, it is interesting to note that the real-life condition of Fat Albert with sugared snack and breakfast commercials inserted throughout the program was still effective in significantly reducing the number of less nutritious foods the children selected.

It is evident that at least this sample of children knew which foods were more nutritious and which were less so. Nevertheless, this appears to have made little difference with regard to the foods they actually selected (something often observed with adults as well). Changes in food choices appear to have developed as a result of what Cosby (1976) calls "affective education" rather than cognitive education. The latter approach is often typified by schools which:

.... often concern themselves solely with the development of cognitive skills, leaving the development of the affective domain to the parents. In doing so what they overlook is the total development of the child (Cosby, 1976, p. 61).

Fat Albert is a human hero who children can empathize with as he struggles with value conflicts and peer group problems. The program themes explore the intangibles of values and ethics, influencing behavior and feelings. It is attitude formation (Cosby 1976, p. 33).

Using Fat Albert or the PSA's to increase the attractiveness and salience of the more nutritious foods seems to have served as a cue that led children to develop more positive attitudes towards these foods and to consider them more actively and positively than they otherwise might have. Similarly the primary purpose for exposing children to snack food commercials in the course of the experiment was to make these foods also more salient. The appropriate directionality of the differences between the control and commercial groups suggests that this goal was accomplished. It may be noted that because the "control" group would typically have been exposed to the same commercials during their normal viewing hours they did not constitute a "control" in the strict sense of the term. Thus, while the groups exposed to food commercials selected a larger number of less nutritious foods than did the control group, it is not surprising to note that this difference was not significant.

The possibility of demand characteristics inducing change in an advocated direction is an everpresent problem in experimental design. In this study, the fact that the Fat Albert program was more effective than PSA's would seem to argue against an explanation of results based on demand characteristics. The experimenters' intentions in each condition should have been equally clear and had both groups responded on the basis of demand characteristics they both should have changed approximately equally, which was not the case. The experimenter also tried to eliminate any momentary set the children may have had after the program by interposing a 5 minute irrelevant discussion between the program and the questions the children were asked. While this diversion could have been longer, it should be noted that children's response sets have generally been found to change particularly rapidly (Mischel, 1972).

Indeed their excitement in expressing their attitudes about their favorite TV programs (the focus of the intervening discussion) strongly suggested that they had moved away from any initial set they may have developed during the TV program.

With regard to demand characteristics, it is also important to note that a number of studies have observed that young children do not understand the motives and intent behind TV commercials (e.g. Robertson & Rossiter, 1974). As a result, they are not likely to be moved to hypothesize what is expected of them in an experiment, as a function of the messages to which they are exposed. Utilizing young children in an experiment, in many regards ensures a relatively "naive" population.

This study considered sugard snack foods and sugared breakfast cereals as belonging together toward one end of the low-high nutritious spectrum. Indeed, at least anecdotal evidence suggests that many children treat sugared breakfast cereals as snacks rather than breakfast foods thus blurring the distinction between the two. Nevertheless, there may be good reason to treat the two types of foods separately. For example, the role of cereal in a balanced breakfast is a complex issue and one that perhaps could be addressed more comprehensively than was the case in this study. In fact, the Fat Albert program contained only one very brief reference to breakfast cereals and only one of the PSA's dealt with breakfast cereals. This suggests that an increased focus on breakfast foods in this study might have resulted in greater differences with regard to breakfast food choices. However, the results provide some evidence counter to this hypothesis. The one PSA that did deal with breakfast cereal was much more explicit with regard to the problems associated with sugared cereals than the Fat Albert program was, yet only the children who viewed Fat Albert selected fewer less nutritious breakfast foods relative to the control group or those seeing Yogi's Gang plus commercials. The relative absence of significance with regard to the breakfast food choices in this study may be more readily understood by examining the dependent measures. The inclusion of only two sets of breakfast foods provided a relatively reduced range of potential responses. Moreover, the repetition of all the alternatives except the cereals further constrained the range of possible responses. In sum, future studies that treat breakfast foods and sugared snacks separately could encourage greater specification and elaboration with regard to both the independent and dependent variables.

The general discussion that followed the administration of the questionnaire suggested that children may well consider the issues raised during the experiment with much more specificity than the global "good-bad" "healthy-non healthy" measure that was obtained. For example, when asked what people should do if they were constantly tempted to "eat things that were not good for them" responses included the following: "they should try and avoid the aisles in the supermarket with all of these foods"; "they should brush their teeth after each snack"; "they should try to get good things like fruit inside the chocolate". Thus when pressed further, children appear to be able to treat the nutritional issue in a more multidimensional manner. In fact, while the global awareness of whether foods are or are not nutritious may not be predictive of actual choice behavior, the more specific cognitions children have, or can be taught to have, may be more useful as predictors in future research.

Footnotes

1. The cooperation of many individuals at the Institute for Communication Research, Stanford University, in

particular Don Roberts, is gratefully acknowledged. The assistance of Doreen Croft, Carol Goldberg and Laurel Ward was also appreciated. Thanks are due to Sally Williams of the Committee on Children's Television, San Francisco; Lou Scheimer and Norman Prescott of Filmation, Reseda CA and Lee Hambleton of CFCF-TV, Montreal for much of the TV materials provided.

2. While Kool-Aid may seem to be an unusual breakfast alternative, in fact, it has been found (e.g. Jenkins et al, 1975) that a small but significant proportion of children do drink Kool-Aid or Pop for breakfast.

References

- Action for Children's Television, Newton, Massachusetts, Petition to Promulgate a Rule Prohibiting the Advertising of Candy to Children on Television, Petition to Federal Trade Commission, April 1977.
- Adler Richard et al, Research on the Effects of Television Advertising on Children: Review and Recommendation. Report prepared for National Science Foundation Research Applications Directorate, 1977.
- Atkin, Charles, "Effects of Television Advertising on Children -- Parent-Child Communication in Supermarket Breakfast Selection," Report #7, Michigan State University, June, 1975a.
- Atkin, Charles, "Effects of Television Advertising on Children -- First Year Experimental Evidence," Report #1, Michigan State University, June, 1975b.
- Clancy-Hepburn, Katherine, Anthony A. Hickey and Gayle Neville, Children's Behavior Responses to TV Food Advertisements, Journal of Nutrition Education, 6 (3), 1974, pp. 93-96.
- Cosby, William, Jr., An Integration of the Visual Media Via Fat Albert and The Cosby Kids into the Elementary School Curriculum as a Teaching Aid and Vehicle to Achieve Increased Learning, Unpublished Ph.D. Thesis, University of Massachusetts, Amherst Mass., 1976.
- Galst, Joann and Mary White, "The Unhealthy Persuader: The Reinforcing Value of Television and Children's Purchase-Influencing Attempts at the Supermarket," Child Development, 1976, 47.
- Goldberg, Marvin E. and Gerald J. Gorn, "Children's Reactions to Television Advertising: An Experimental Approach," Journal of Consumer Research, Vol. 1, Sept. 1974, pp. 69-75.
- Goldberg, Marvin E. and Gerald J. Gorn, "Children's TV Commercials: Do Child Viewers Become Satiated Too?" Working Paper, Faculty of Management, McGill University, Montreal.
- Gorn, Gerald J. and Marvin E. Goldberg, "The Impact of Television Advertising on Children from Low Income Families," Journal of Consumer Research, In Press, September, 1977.
- Greenberg, Allan and Charles Suttoni, "Television Commercial Wearout," Journal of Advertising Research, Vol. 13, No. 5, October, 1973, p.47-54.
- Gussow, Joan, "Counternutritional Messages of TV Ads Aimed at Children," Journal of Nutrition, 4 (1972), pp. 48-52.
- Jenkins, Suzanne, Mary Stumo and Jane Voichick, "Evaluation of the Nutrition Film Series 'Mulligan Stew'", Journal of Nutrition Education, 7 (1), 1975 p. 17-19.
- Mauro, Frank and Feins, R.P. "Kids, Food and Television: The Compelling Case for State Action", Office of Research and Analysis Program and Committee Staff. New York State Assembly, Albany N.Y., March 1977.
- Mischel, Walter, Ebbe B. Ebbesen and Antonette P. Zeiss, "Cognitive and Attentional Mechanisms in Delay of Gratification," Journal of Personality and Social Psychology, 21 (2), 1972, pp. 204-218.
- National Association of Broadcasters, The Television Code, 18th edition, June 1975.
- Poulos, Rita Wicks, "Unintentional Negative Effects of Food Commercials on Children: A Case Study," Media Action Research Center, 1975.
- Robertson, Thomas S. and John R. Rossiter, "Children and Commercial Persuasion: An Attribution Theory Analysis," Journal of Consumer Research, 1 (June 1974), pp. 13-20.
- Rubin, Ronald S., "An Exploratory Investigation of Children's Responses to Commercial Content of Television Advertising in Relation to their Stages of Cognitive Development," unpublished doctoral dissertation, University of Massachusetts, Amherst, Mass., 1972.
- Society for Nutrition Education (ed.), Audiovisuals for Nutrition Education, Nutrition Education Resource Series, No. 9, Society for Nutrition Education, Berkeley, California.
- Ward, Scott and Daniel Wackman, "Family and Media Influences on Adolescent Consumer Learning," American Behavioral Scientist, 14 (3) January 1971.
- White House Conference on Food, Nutrition and Health Final Report, Washington, D.C., Government Printing Office, 1971.

PSYCHOMETRIC CHARACTERISTICS OF BEHAVIORAL PROCESS DATA:
PRELIMINARY FINDINGS ON VALIDITY AND RELIABILITY

Jacob Jacoby, Purdue University
Robert W. Chestnut, Columbia University
Wayne D. Hoyer (student), Purdue University
David A. Sheluga (student), Purdue University
Michael J. Donahue (student), Purdue University

Abstract

Along with verbal protocols and eye movement techniques, a family of behavioral process methods have recently been introduced to study pre-decision information acquisition and usage. Various questions may be raised in regard to these methods, including those relating to validity and reliability. Data are presented here regarding the validity and reliability of behavioral process methods. A fuller report will be provided in Jacoby and Chestnut (in preparation).

Introduction

As noted elsewhere (Jacoby, 1976, p. 5), all that is usually required for a measure or measurement approach to gain acceptance in that consumer research community is for an investigator to describe and use the measure in a single published study. As long as the measure or approach has a certain degree of face validity -- and oftentimes, even if it does not -- it tends to be uncritically accepted.

This uncritical attitude has led to a profusion of operational measures for virtually all of our core constructs. As an example, Jacoby and Chestnut (in press) identify 55 different brand loyalty indices in the published literature and came across several others in the unpublished proprietary literature. The problem, as Kohn and Jacoby (1973) demonstrated in regard to operationally defining innovators, is that different measures of the same construct, when applied to the same subject population, often produce conflicting results. Under these circumstances -- namely, a welter of different definitions which yield contradictory results -- it becomes exceedingly difficult to compare and integrate findings from different investigations. The end result is that consumer research is being strangled by its poor measures and progress in understanding consumer behavior is thwarted. Clearly, we must begin to devote greater empirical attention to describing the psychometric properties of our measures and measurement approaches, particularly the validity and reliability of said measures and approaches. This becomes especially true where we have the possibility of rejecting such measures before they can establish a toehold and proliferate.

Having advocated this in general (Jacoby, 1976; Jacoby and Chestnut, in press), it would be hypocritical if we did not apply this logic in our own research. In particular, since 1972-73 (see Jacoby, 1975a) we have been involved in the development and extension of an empirical approach for the purposes of identifying pre-decision information acquisition. An overview of approximately 20 studies in this series is described in Jacoby (1977) and theoretical underpinnings have been outlined in Chestnut and Jacoby (1977a). More recent developments include: several papers published or in press (Jacoby, Chestnut, and Fisher; Jacoby, Chestnut, and Silberman; Sheluga and Jacoby); completed studies on concept learning (Donahue and Jacoby, in preparation), attribution theory (Major and Chestnut, in preparation), and salesman influence in life insurance purchases

(Chestnut, 1977); an illustration of its application to health care services (Chestnut and Jacoby, 1977b); the completion of a monograph on consumer nondurable purchases (Jacoby and Chestnut, 1977); and an on-going project for the Federal Trade Commission (Jacoby and Chestnut, in progress). Much of this work is now being integrated into a single volume (Jacoby and Chestnut, in preparation).¹

The approach has also stimulated acceptance and usage by others. Examples from academia include: Bettman (e.g., Bettman and Kakkar, 1977), who was introduced to the approach during a 1975 visit to Purdue and through his participation as a consultant (Bettman and Jacoby, 1976) on the Jacoby and Chestnut (1977; Jacoby, Chestnut, Weigl, and Fisher, 1976) project; Van Raaij (1977a,b), who became acquainted with the approach through a colloquium (Jacoby, 1975b) at the University of Tilburg in May of 1975; and investigators in Mannheim, West Germany where the senior author spent the summers of 1975 and 1976 as a visiting scholar (see Raffée et al., 1976). Further, a variety of commercial applications emanating from this research program have been developed and refined, and are now being offered by the Opinion Research Corporation of Princeton, New Jersey.

Given the burgeoning interest and activity in this approach, it is incumbent upon those who utilize it to begin supplying data regarding its psychometric properties. The present report focuses on validity and reliability, bringing together some unpublished data from several investigations which bear on these issues.²

Validity

Addressing validity requires an understanding of the basic kinds of behavioral process (BP) data generated and how these relate to the criteria available. Reduced to its essential core, the BP approach provides data regarding (1) information acquisition behavior -- the information acquired prior to arriving at a decision (which may be further decomposed into the depth, content, and sequence of information acquisition), and (2) choice behavior -- the identity of the option selected when the decision is reached. Validity

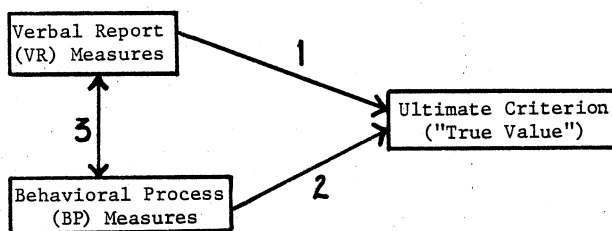
¹An approach similar to ours, was independently arrived at by Payne (e.g., 1976).

²The data on a third psychometric property, sensitivity, will not be detailed in this report. Suffice it to say that BP is believed to provide much greater sensitivity, especially in regard to the sequence of information accessing and usage. Particularly in a crowded product category (where there may be 20 or more different purchase options and 30 or more items of objective information for each), it would seem beyond the capabilities of most people to be able to accurately recall the precise amount, identity, and exact sequence in which information was accessed and used.

assessment in these two cases assumes different forms because of the nature of the criteria available.

In the case of information acquisition behavior, there is no external "true value" criterion to serve as a standard and against which we can compare our data. That is, we have no way of crawling into the person's head and directly measuring the information that he acquires in vivo when arriving at a decision in the "real world." Thus, regardless of whether we solicit data using traditional verbal reports (via questionnaire or interview, e.g., "Tell me what information you are acquiring and which information you are ignoring when you make your breakfast cereal purchase decision"), employ verbal protocols (e.g., Bettman, 1970), use a behavioral process approach, or employ an eye tracking procedure (e.g., Russo and Rosen, 1975), we have no ultimate criterion (or "true value") against which to evaluate our data. This is reflected in Figure 1 which depicts verbal reports (VR) and behavioral process (BP) as alternative approaches for assessing actual pre-decision

FIGURE 1
Relationships Existing Between Verbal Report and Behavioral Process Measures of Behavior (Either Information Acquisition/Usage or Choice) and Actual In Vivo Behavior



information acquisition. Since we have no way to determine actual information acquisition during real world decision-making, relationships 1 and 2 are not ascertainable. What is measurable is the relationship between the verbal report and BP approach, Relationship 3. This is termed "convergent validity" and represents the necessary first step for demonstrating construct validity (see Campbell and Fiske, 1959; Nunnally, 1967).³

In contrast, in the case of choice behavior, both verbal reports of choice (e.g., "I usually buy Rice Krispies") and the identity of the option selected at the end of a BP simulation are potentially verifiable against an external criterion. That is, if Mrs. Jones says she usually buys Rice Krispies (in response to soliciting a VR), or ends up selecting that brand in a BP simulation, then (at least in theory) it is possible to verify this fact. Hence, all three relationships depicted in Figure 1 are potentially ascertainable for choice behavior.

Information Acquisition Behavior

Convergent validity represents the degree to which different operational measures of the same concept yield comparable results (Relationship 3 in Figure 1), i.e., "a confirmation by independent measurement procedures" (Campbell and Fiske, 1959, p. 81). Data on the convergence between VR and BP measures of information usage are available from several investigations. The earliest example comes from a study conducted during the spring of 1973. In this investigation, Jacoby, Szybillo, and Busato-Schach (1977, p. 213) report a median correlation

³The various types of validity and their relationship to each other are discussed in Jacoby and Chestnut (in press, Chapter 4).

of .37 (between verbal reports of typical information usage behavior and the order in which these types of information were accessed in a BP simulation) for students reaching a brand choice decision under a Brand Names Present (BNP) condition, and .53 for students arriving at choices under a Brand Names Absent (BNA) condition.⁴ A replication conducted two years later in Germany (Raffée et al., 1976) reveals slightly higher coefficients when VRs of typical usage behavior were correlated with actual usage/non-usage in the BP task: a median of .63 in the BNA condition and .61 in the BNP condition (see Jacoby, Hefner, Schöler, Grabicke, and Raffée, 1976, p. 9).

Probably the most comprehensive evidence on convergent validity was obtained in an investigation (Jacoby and Chestnut, 1977)⁵ which employed sociodemographically heterogeneous samples of residents from Tippecanoe County, Indiana, three different test products, and a much more advanced and sensitive BP methodology than was used in the two above-cited investigations. Table 1 presents the data for the information dimensions used by

TABLE 1
Relationship of Self Reports of Typical Accessing Behavior with Actual Behavior Displayed in Decision Task

	BNP	BNA
Breakfast Cereal Dimensions	N = 127	N = 0
Addr. for Cooking Info.	-.05	
Addr. for More Nutr. Info.	.07	
Basic Ingredients	.37***	
Calories	.20*	
Carbohydrate Content	.43***	
Cooking Hints	.19*	
Comparative Cost Info.	.12	
Coupons	.37***	
Diet Info.	.30***	
Expiration	.35***	
Fat Content	.23**	
Items for Children	.35***	
Gen. Nutr. Info.	.20*	
Guarantee	.04	
Background on Brand	-.01	
Kosher	---	
Manfr's Name & Address	.21**	
Net Weight	.29***	
Offers to Adults	.15*	
Offers to Children	.32***	
Preservatives	.39***	
% U.S. RDA	.07	
Price	.34***	
Protein Content	.08	
Recipes	.25**	
Servings per Box	.04	
Sug. Serving Size	.22**	
Unit Price	.22**	
Vitamin Content (% MDR)	.33***	
Vitamin Supplements	.39***	
Median Spearman:	.22	

⁴The Brand Names Absent (BNA) manipulation is used to reduce the impact of relevant information in memory on information acquisition behavior. In essence, it makes the BP simulation comparable to a situation in which the subject is confronted with a set of purchase alternatives consisting entirely of "new" brands.

⁵This study, hereafter referred to as the NSF investigation, was supported by a grant from the National Science Foundation (GI-43687).

TABLE 1 (Cont'd)

	BNP	BNA
Margarine Dimensions	N = 68	N = 110
Addr. for Nutr. Info.	-.15	.13
Basic Ingredients	.55***	.42***
Expiration Date	.59***	.41***
Calories per Serving	.32**	.42***
Carbohydrate Content	.26*	.29**
Cooking Hints	.35**	.46***
Coupons	.59***	.37***
Fat Content	.16	.20*
Guarantee	.02	.34***
Kosher	-.13	.03
Mfr's Name & Address	-.07	.50***
Net Weight	.23*	.29**
Parve	-.09	-.21*
Price	.22*	.24**
Product Claims	.25*	.20*
Protein Content	.12	.17*
Servings per Pound	.28**	.22**
Special Offers	.00	.17*
Storage Instructions	-.35**	.33***
Suggested Serving Size	---	.02
Type/No. Package Units	.35**	.36***
Unit Price	.10	.26**
Vitamin Content	.33**	.35**
Median Spearman:	.23	.29
Headache Remedy Dimensions	N = 56	N = 95
Active Ingredients	.29*	.60***
Addr. Info. from Mfr.	.32**	.32***
Dosage	.38**	.27**
Expiration Date	.31*	.08
Health Warning	.48***	.40***
Container Safety Instructions	.08	.14
Mfr's Name & Address	-.06	.35***
No. of Tablets per Bottle	.39**	.28**
Price	.45***	.40***
Type of Cap	.31*	.37***
Unit Price	.21	.23*
Use for these Symptoms	.57***	.40***
Warning about Children	.37**	.33***
Median Spearman:	.32	.33

* = $p < .05$; ** = $p < .01$; *** = $p < .001$

the 456 (out of the 615) subjects in this study who were information seekers. (N.B. The other 149 subjects made their brand choices on the basis of seeing only the brand names of the 16 available brands and did not acquire any additional information prior to arriving at their decision.) As can be seen, the median Spearman coefficients (between self reports of typical information accessing behavior and behavior displayed in the BP decision task) across all information dimensions in the five samples ranged from a low of .22 to a high of .33. Again, values were higher for the BNA than for the BNP condition. Since analysis of the depth of search for these subjects (see Table 8 discussed below) revealed that the mean number of different information dimensions considered prior to making their purchase decisions was 3 to 4 in the BNP condition and approximately 7 in the BNA condition, median Spearman coefficients were identified when only the five most heavily considered dimensions for each of the five samples were considered. These coefficients are: .29 for breakfast cereal; .55 for margarine, BNP; .37 for margarine, BNA; .45 for headache remedies, BNP; and .40 for headache remedies, BNA. Thus, convergent validity seems to be higher (near .40) if attention is limited to frequently used dimensions. Also of interest, for the first time, BNP coefficients are higher than their BNA counterparts.

In sum, three findings are noteworthy in regard to the convergent validity between VR and BP measures of information acquisition/usage. (1) In general, it would appear that the information people say they typically acquire and the information they actually do acquire in a behavioral process simulation tends to be positively related, but only to a small degree. Coefficients of Determinance (r^2) reveals that only between 5% to 40% of the variance is explained. If we omit the two samples using an earlier and less sensitive BP variant of methodology, the remaining Coefficients of Determinance based on full samples and all available information dimensions all fall below 12%. (2) Higher correlation coefficients are obtained when brand names are absent relative to when they are present. This makes considerable sense since, when brand names are present, the individual may be accessing information from memory as well as from the external information environment. Unfortunately, BP procedures can only directly assess information acquired from the external environment. When brand names are absent, all brand-specific information must be acquired from the environment. Thus, without memory to rely on, there should be a greater correspondence between what people say is important to them and what they pay attention to (i.e., access) in a BP simulation. (3) Higher coefficients are obtained for the more heavily accessed information dimensions. Again, this makes considerable sense: people can be expected to less accurately recall accessing infrequently used rather than frequently used information -- especially for low-priced nondurables where much of the information is of a relatively trivial nature.

Several additional comments should be made in regard to the low ($r = .2$ to $.3$) correlations obtained between VR and BP indicants when all available information is included and the newer methodological variant is used. First, these coefficients could have been negative, in which case we would have had a major dilemma on our hands (i.e., since there is no available external criterion, which of the two measures do we say provides the "reasonably correct" picture?). On the other hand, these two approaches could have yielded indicants that were in very high agreement (say, $r = .8$ or $.9$). In such a case, it could be argued that the BP measures do not provide much unique coverage and (given that they are more difficult, expensive, time-consuming, etc., to collect) should therefore be dispensed with. From this perspective, coefficients in a low to mid positive range (i.e., $.3$ to $.7$) would seem to be optimal.

Even granting this argument, the obtained coefficients still show a relatively low degree of relationship and suggest that the measures do contain a fair share of error. However, these low coefficients also seem to be at least partially attributable to factors above and beyond unexplained error. In responding to the VR measure, subjects may have indicated what their behavior may have been like at an earlier point in time but, perhaps (aided by memory), they no longer do. Further, the VR measure asked respondents to indicate what they do "in general," i.e., across purchase instances. In contrast, the BP situation examines only a single instance. As argued elsewhere (Jacoby and Chestnut, in press) in regard to brand loyalty, if we are trying to describe a person's customary behavior, then sampling considerations suggest that we need to consider his behavior across several rather than a single test instance. Stated in a somewhat different form, the two assessments (i.e., VR and BP) were measuring slightly different things: the VR measure asked the individual to indicate what he did in general, whereas the BP measure considered only a single instance. In a similar vein, the BNA condition and the VR data also address somewhat different things. Whereas both the VR measures and BP-BNP condition attempt to reflect and examine the external information environment as it exists in the real world,

the BNA condition, through the very absence of familiar brand names, does not. Again, to the extent that different things were being addressed by the two measures (i.e., VR and BP-BNA), we would expect this to be reflected in a lowered correlation. Thus, considering the various actual and potential influences on these coefficients, the values obtained seem quite reasonable and in line with what might have been expected.

Choice Behavior

Both BP and VR methods may be used to collect choice data. Important questions are: (1) just how much agreement exists between these two approaches (Relationship 3 in Figure 1), and (2) how accurate are BP-derived estimates of actual choice behavior (Relationship 2) relative to VR-derived estimates of this same behavior (Relationship 1)? Respectively, these are questions of convergent and criterion-related validity.

Data regarding both types of validity comes from the three BNP samples in the previously noted NSF investigation (Jacoby and Chestnut, 1977). In a fashion similar to numerous other investigations, each subject was asked to recall the names of all brands in the product category he had purchased during the preceeding 12 months, and to then estimate the percent of purchases he devoted to each of these brands. The sum of the estimated percent of purchase for each brand was then divided by the total *n* in each sample to arrive at an aggregated VR estimate of each brand's market share for this sample. Since brands other than those included in our simulation were recalled (e.g., according to Time magazine [1976, p. 61]): "there are no fewer than 156 brands" of breakfast cereal available in American supermarkets and our study considered only 16 of those with the largest market shares), the VR market share estimates cumulated across all brands in our investigation sum to less than 100%. The proportion of time each brand ended up being "purchased" at the end of the BP simulation task represented that brand's BP-derived market share.

Since BP estimates were based upon only those 16 brands provided in each of the decision simulations, these estimates summed to 100% for each product category. In contrast, the VR market shares for 198 subjects across the 16 brands of breakfast cereal used in the investigation cumulated to 69.3%; the 16 brands of margarine had a combined VR market share (across *n* = 107) of 52.3%; and the 16 brands of headache remedy had a combined VR market share (across *n* = 104) of 90.6% (see Table 2). Let us consider the results for each of these products while bearing this point in mind.

TABLE 2
Unadjusted Market Share Estimates

Breakfast Cereals	BP	VR
Post Raisin Bran	15.17%	3.28%
Quaker 100% Natural	14.69	5.93
Shredded Wheat	10.90	6.74
Corn Flakes	8.53	7.82
Special K	7.11	5.20
Grapenuts	6.64	4.35
Rice Krispies	6.16	9.08
Total	6.16	3.41
Sugar Frosted Flakes	5.21	3.36
Kellogg's Raisin Bran	5.21	2.97
Cheerios	4.27	6.72
Heartland Natural	3.79	2.51
Wheaties	2.37	4.03
Super Sugar Crisp	1.90	2.03
Fruit Loops	.95	1.57
Cap n'Crunch	.95	.30

Σ = 100.00% 69.30%

TABLE 2 (Cont'd)

Margarine	BP	VR
Blue Bonnet	23.01%	17.17%
Fleishmann's Soft	13.27	1.63
Imperial	10.62	10.74
Mazola	7.96	6.45
Nu-Maid	7.96	4.59
Chiffon	6.19	1.18
Soft Blue Bonnet	6.19	.00
Diet Parkay	4.42	2.30
Bluebrook	4.42	.10
Mrs. Filberts	3.54	1.94
Hillfarm	3.54	.13
Land O'Lakes	2.65	1.47
Promise	2.65	.87
Whipped Miracle	1.77	1.22
Squeeze Parkay	.88	1.43
Whipped Blue Bonnet	.88	1.03

Σ = 100.00% 52.25%

Headache Remedies	BP	VR
Bayer	23.08%	23.56%
Bufferin	17.31	19.31
Tylenol	14.42	12.33
Anacin	11.54	12.87
Excedrin	8.65	8.42
Lily A.S.A.	3.85	4.16
Vanquish	3.85	1.19
Arthritis Pain Formula	3.85	1.19
Arthritis Bufferin	2.88	.50
Osco	1.92	3.96
Norwich	1.92	2.57
Osco APC	1.92	.99
Bayer Timed Release	1.92	.69
Empirin	1.92	.00
Excedrin PM	.96	.00
Phenodyne Bluline	.00	.00

Σ = 100.00% 90.95%

The Pearson *r* calculated between the BP and VR market share estimates for breakfast cereal was .48. Two factors are relevant to an evaluation of this coefficient. First, due to an oversight, when subjects indicated that they had purchased Raisin Bran during the preceding 12 months, the interviewers had not been instructed to probe with a follow-up question in order to determine just which brand (Kellogg's or Post) was involved. Hence, the VR data for these subjects could not be counted for either of these brands and were therefore omitted from the calculation of the VR market shares. (The BP data suggest that most of these purchasers would have specified Post.) Second, Quaker 100% Natural had recently been introduced into the market area and had not had sufficient time to have been purchased with any great regularity during the preceding year. There is thus reason to believe that the VR estimates of market share for the year were not an accurate reflection of purchase probabilities existing at the time of the investigation. When these 3 problem brands are removed from the set of 16 test brands, the Pearson coefficient increases to .73 (*p* < .01). With only the two Raisin Bran's removed, the coefficient increases to .64.

The Pearson coefficient for margarine is a highly significant .83. However, a problem exists similar to that for breakfast cereal. If a respondent recalled Blue Bonnet but did not differentiate just which Blue Bonnet was meant, these data could not be used in calculating the VR market share estimates. With the three Blue Bonnet brands omitted, the Pearson Coefficient decreased to .57 (*p* < .05).

The VR and BP estimates for headache remedies seem to be in fairly high agreement, and a Pearson product moment

correlation between the two was highly significant ($r = .98$). The high (90.6%) percentage of VR market share estimates entering into this correlation probably makes it the most dependable one of the three samples.

It should be acknowledged, however, that the obtained correlations, while moderately high (see Table 3), are based on only 16 (or 13) sets of figures so that changes in 2 or 3 pairs of figures would be sufficient to exert dramatic impact on the coefficients obtained. Second,

TABLE 3
Correlations between VR-Derived and BP-Derived Market Share Estimates

Product	Number of Brands	Cumulative VR Market Share	r	p
Breakfast Cereal (N = 198)	All 16	69.3%	.48	$\approx .06$
	14 (-2 Raisin Bran)	63.0%	.64	$< .02$
	13 (-2 RB & Quaker 100% Natural)*	57.1%	.73	$< .02$
Margarine (n = 107)	All 16	52.3%	.83	$< .001$
	13 (-3 Blue Bonnet)*	34.1%	.57	$< .05$
Headache Remedies (N = 104)	All 16*	91.0%	.98	$< .001$

* = Adjusted for confounds (i.e., considered to provide the best estimate).

we are again confronted with the problem that the VR measure assessed behavior over a series of purchase instances whereas the BP indicant was based on only a single instance. Finally, it should also be noted from Table 3 that the cumulated VR market share estimates (when adjusted for confounds) for a given product appear to be positively related to the strength of the relationship manifested between VR and BP estimates -- the greater the amount of VR market share included in our BP test situation, the higher the relationship between the two estimates.

The NSF investigation also supplied relatively crude data regarding Relationship 2 in Figure 1. Readers familiar with this research from earlier reports (e.g., Jacoby, 1975a; Jacoby, Chestnut, Fisher, and Weigl, 1976; Bettman and Jacoby, 1976) may recall that subjects were provided with "cents-off" coupons which they were later able to apply toward purchases of their test product. Breakfast cereal and margarine subjects received five such "30¢-off" coupons, with three of these being good only toward purchases of the brand they had selected at the end of their BP choice task. The other two coupons were good toward the purchase of any brand in that product category, be it the one they had selected in the BP simulation or another. The coupons were valid for a period of four months from the date that the individual participated in the study.

An indication of predictive validity is obtained by addressing the following question: To what extent was the brand selected in the BP task predictive of the brand later purchased (Relationship 2, Figure 1)? Table 4 provides data pertaining to predictive validity for the two margarine conditions. While the vast majority of people redeemed at least one coupon (90% in the BNP sample and 83% in the BNA sample), considerably more coupons were redeemed for the BNP sample (64%) than for the BNA sample (51%). (The data for BNP are comparable to those obtained with breakfast cereal, where 89% of the subjects redeemed at least one coupon and 67% of all coupons were returned.) This pattern of coupon

TABLE 4
Coupon Redemptions for Margarine

	BNs Present			BNs Absent		
	Pos- sible	Actual	%	Pos- sible	Actual	%
Subjects redeeming coupons	107	96	90%	110	91	83%
Designated coupons redeemed	321	216	67%	330	170	52%
Undesignated coupons redeemed	214	128	60%	220	113	51%
-BP selected brand	214	52	24%	220	12	5%
-other 15 brands	214	76	36%	220	101	46%
Total coupons redeemed	535	344	64%	550	283	51%

redemptions suggests that the coupons had value for the subjects, and providing five such coupons most probably generated a realistic degree of task motivation, as was intended. Unfortunately, using designated coupons as a manipulation of motivation meant that validity data based on these coupons would be confounded. Moreover, while redemption of the undesignated coupons provides somewhat cleaner data, they are nonetheless affected by the fact that each subject also possessed three designated coupons and these might have influenced how he used his undesignated coupons.

Notwithstanding these problems, Table 4 reveals the following pattern: (1) A higher proportion of coupons were returned by the BNP (as compared to the BNA) subjects, both overall and for the brand they selected in the BP task. This is not surprising since 72% of BNA subjects chose a brand at the end of the BP task which they supposedly didn't purchase (according to an independent VR assessment). The motivation for these subjects to redeem their coupons could be assumed to be low. (2) In contrast, BNA subjects used a higher proportion of their undesignated coupons to redeem unselected brands. (3) Approximately 10% of the undesignated coupons returned by the BNA sample was for their selected brand, while this figure was 40% (24/60) for the BNP sample. The 10% figure for the BNA sample is only slightly higher than the 6.125% that could be expected to have selected their favorite brand in the BP choice task (i.e., there being 1 out of 16 chances of selecting this brand). The 40% figure for the BNP sample is just about what one might have expected given that these subjects already possessed three coupons valid only for their selected brands (and might have therefore chosen one other brand for the sake of variety) and also because the weighted mean VR brand loyalty score for this group was 48 on a 100 point base, indicating that slightly less than half their purchases were devoted to their most often purchased brand in that product category.

Reliability

The key theme of reliability is that of consistency in response. Such consistency can be examined at various levels. As applied according to the traditional approach, an attempt is made to hold virtually everything constant in regard to the assessment except the fact that assessment takes place at two different points in time. That is, the same measuring device (or approach) is applied to the same individuals, in regard to the same stimuli, in the same testing environment, by the same test administrator, etc. However, to be able to generalize, one must also explore consistency in response across different subjects from the universe of

all possible subjects of concern, different stimuli from the universe of all possible stimuli of concern, different testers, different environments, etc. That is, according to an increasingly influential contemporary interpretation, traditional reliability assessment expands into the quest for generalizability (see Cronbach, Gleser, Nanda, and Rajaratnam, 1972).

This paper describes data relating to several points along this consistency continuum. First examined is the consistency of response within the same subject across two different points in time with respect to the same decision. Next, we consider the consistency of response within the same subject across qualitatively different decisions. Finally, we consider the consistency of response across different subjects and different decisions.

Further, while the section on validity considered both information acquisition and choice behavior, hardly any relevant data have been collected regarding the reliability of choice. Hence, our attention is generally confined to pre-decision information acquisition, particularly the depth (i.e., how much?) and sequence (in what order?) of such information acquisition.

Same Subject, Same Decision, Cross-Time Consistency

One fundamental form of traditional reliability is test-retest reliability -- that is, will the same measurement procedures administered at two different points in time to the same individuals and under the same conditions yield identical (or nearly identical) results? Naturally, this assumes no relevant change during the period intervening between the first and second assessment in the item(s) being assessed. Another basic variety of traditional reliability is parallel form -- that is, to what extent do alternative or equivalent forms of an instrument (or, in this case, a measurement procedure), when applied to the same content, yield comparable results?

Preliminary answers to these questions come from a study in which 71 undergraduates were asked to arrive at a "blind date" choice from among 8 hypothetical options, each described along 15 information dimensions. Approximately half this sample was re-tested after a ten-minute delay during which they were engaged in an interpolated activity; the other half was re-tested approximately one week later. Further, some subjects were given the identical information, with only the labels designating the alternatives being changed (either from letters to numbers or vice versa), while other subjects were confronted with entirely new information for eight different blind dates. The Different condition reflected parallel forms reliability, while the Same condition represented test-retest reliability.

Table 5 indicates the number of subjects in each condi-

TABLE 5
Within Subject, Within Decision (Test-Retest and Parallel Form) Consistency

	Immediate		Delayed	
	Same (N=17)	Different (N=25)	Same (N=17)	Different (N=12)
No. of items acquired	.67***	.91***	.27	.69**
No. of "dates" consdrd.	.63**	.87***	.01	.53*
No. of dimens. consdrd.	.50*	.90***	.53*	.46
Type 2 transitions	.64**	.79***	.09	.58*
Type 3 transitions	.69***	.89***	.45*	.85***

tion and correlations obtained. In general, these coef-

ficients reveal a pattern of high test-retest reliability for the Immediate retest condition, particularly when Different rather than Same stimuli were employed. Also, and as might have been expected, coefficients were generally lower under the Delayed condition. However, they are quite comparable to the test-retest reliabilities obtained for a set of VR items collected for comparison purposes from the these Delayed retest subjects (see Table 6). These latter data suggest that the Delayed/

TABLE 6
Consistency in VR Responses for Subjects in the Delayed Retest Condition

	Same	Different
1. What is your current status with respect to members of the opposite sex? (check one)	.73***	.50*
<input type="checkbox"/> Married		
<input type="checkbox"/> Formally engaged		
<input type="checkbox"/> Going steady with one person		
<input type="checkbox"/> Currently dating two or more people on a regular basis		
<input type="checkbox"/> Not currently dating anyone		
<input type="checkbox"/> Never had a date with a member of the opposite sex		
<input type="checkbox"/> Other (please specify) _____		
2. For your answer in question 1, how long has this been the case? _____ months.	.26	.71*
3. How often do you date? Approximately _____ times per month.	.47*	.55*
4. Have you <u>ever</u> had a "Blind date?" <input type="checkbox"/> Yes <input type="checkbox"/> No	.01	.58*
5. If yes, approximately how many times? _____	.90***	.86***
6. (Pre-Task) Even if you have never had a blind date, make believe you were going to go out on a date with someone you never met before. What would you say the chances would be that you would have a good time? (Response options consisted of a modified 11-point Juster scale; see Engel, Kollat, and Blackwell, 1973, p. 363).	.29	.81***
7. (Post-Task) From the information you have now, what probability would you say there was that you would have a good time if you went out with this person? (modified 11-point Juster scale)	.14	.71**

Same subjects might have been somewhat strange. In particular, consider the coefficients for items 2 and 4. Additional data are being collected to clarify this issue.

The two Same cells provide our only indication of the consistency of choice (in contrast to the consistency of information acquisition behavior, which we have been considering up to this point). For any given blind date selected at t_1 , the probability of selecting this same date at t_2 was 1 out of 8 (or .125). Twelve out of 17 (71%) subjects selected the same date on the Immediate retest, while 5 out of 12 (42%) did so for the Delayed retest.

Consistency within the Same Subject across Different Decisions

Thus far, our published reports have considered pre-decision information acquisition for individuals making purchase decisions. However, as suggested by the study above, the BP approach is applicable to other types of decisions as well. One attempt to explore the psychometric properties of the approach involved having the same subjects arrive at three qualitatively different kinds of decisions: purchase one of eight brands of toothpaste; select one of eight (hypothetical) individuals as a blind date; and, choose (vote for) one of the eight candidates on the 1976 U.S. Presidential primary ballot in Indiana. The number of information dimensions for each task was 17, 15, and 6, respectively.

Attention is confined here to the (Brand) Names Absent conditions for the following reasons. First, while the matrix of information used for the toothpaste and political candidate decisions was "real" information (in the sense of actually being available in the environment), this was the only condition possible in the case of the blind dates, where the different alternatives were all hypothetical. Second, when names are present, respondents can rely upon the information they have in memory regarding each of these different options. As a result, we are unable to determine just which information is accessed and considered relevant (since information may also be retrieved from long term memory). Further, since each subject probably had different past experiences with respect to the various toothpastes and candidates, a BNP condition would introduce too many uncontrolled and potentially confounding factors. Finally, when the identity of the choice options is known (as in BNA), a large proportion of subjects (e.g., 36% to 46% in the NSF investigation, see Table 7) do not access any information from the external environment. Thus, (Brand) Names Absent was used to insure that subjects engaged in at least some information acquisition behavior.

Table 8 provides the relevant data for the 31 subjects on 3 depth and 2 sequence variables. All 15 Pearson co-

efficients were positive, highly related (median $r = .57$), and statistically significant -- 14 at beyond the .01 level and 11 of these at beyond the .001 level. Thus, there seems to be a high degree of consistency in the depth and sequence of information acquisition within subjects who arrive at qualitatively different types of decisions.

TABLE 8
Within-Subject, Cross-Decision Consistency

	Candidates	Toothpaste
No. of Items Acquired		
Blind Date	.70***	.41**
Candidates		.72***
No. of Dimensions Considered		
Blind Date	.56***	.50**
Candidates		.42**
Start-to-Finish Decision Time		
Blind Date	.55***	.39*
Candidates		.60***
Type 2 Transitions		
Blind Date	.72***	.61***
Candidates		.57***
Type 3 Transitions		
Blind Date	.76***	.59***
Candidates		.56***

* = $p < .05$; ** = $p < .01$; *** = $p < .001$

Consistency across Different Subjects and Different (but comparable) Decisions

Finally, the NSF investigation (Jacoby and Chestnut, 1977) provides data regarding consistency across heterogeneous decision tasks. The 615 subjects in the 5 samples were interviewed over a 2-year period (1975-76) by more than 12 different interviewers. Respondents reflected a cross-section of Tippecanoe Co. (Indiana) residents above the age of 18 and were representative in terms of age, education, and total family income. All had been "qualified" in advance as users of all three

TABLE 7
Cross-Subject, Cross-Decision Task Consistency

	Brand Names Present			Brand Names Absent	
	B.C.	M	H.R.	M	H.R.
Matrix Size (No. of Brands x No. of Information Dimensions)	16 x 30	16 x 23	16 x 13	16 x 23	16 x 13
Sample Size (N =)	198	107	104	110	96
Percent of Sample that Acquired at Least One Item of Information	64%	64%	54%	100%	100%
DEPTH OF SEARCH:					
Mean no. of items acquired by each subject	8.70	9.10	9.98	29.59	30.58
Median no. of items acquired by each subject	6	6	6	26	26
Mean no. of brands considered by each subject	3.16	4.59	3.39	11.03	12.12
Mean no. of dimensions considered by each subject	4.55	3.37	3.73	7.45	6.73
Mean start-to-finish decision time (in seconds)	140.16	141.32	163.88	374.86	351.02
Mean time devoted to each item (in seconds)	16.11	15.53	16.42	12.26	11.86
CONTENT OF SEARCH:					
Price - % searchers accessing dimens. at least once	78%	72%	66%	76%	60%
Unit Price - " " " " " " " "	25%	15%	29%	41%	45%
Expiration Date - " " " " " " " "	23%	26%	25%	42%	38%
Price - Mean prop. of ea. individual's search of...	.31	.38	.22	.15	.08
Unit Price - " " " " " " " "	.06	.05	.08	.06	.07
Expiration Date - " " " " " " " "	.05	.06	.07	.06	.04
SEQUENCE OF SEARCH:					
Type 1 transitions	.01	.02	.00	.01	.00
Type 2 transitions	.39	.29	.47	.31	.29
Type 3 transitions	.40	.16	.16	.24	.25
Type 4 transitions	.20	.16	.16	.24	.25

test products.

Table 7 presented the relevant data for six depth, six content, and four sequence variables. The depth variables are fundamental and need no further explanation. The content variables cover the three information dimensions common to all three test products. The sequence variables are those which we developed in 1974 and have been using ever since. Readers are directed to Jacoby, Chestnut, Weigl, and Fisher (1976) for a more detailed description of these indices. Given the many possible configuration of values that the data could assume, inspection of Table 7 reveals an exceptionally high degree of consistency, particularly when one considers the three groups nested under BNP as one comparison set, and the two groups under BNA as another.

Discussion

Appearances (i.e., the many tables of data presented) to the contrary, we have here only touched the tip of the data iceberg. Space limitations do not permit a more exhaustive presentation of data or discussion of findings. Based upon what we have been able to present, however, the tentative findings may be summarized as follows:

1. The convergent validity between VR and BP assessments tend to be in the order of .2 to .4 for information acquisition behavior and range from .48 to .98 for choice behavior. One would expect such greater agreement in regard to a single global act such as choice relative to agreement on the more molecular behavior of information usage. One way to look at this finding is to say that the BP approach has more unique variance to contribute in the study of information acquisition behavior.
2. With respect to reliability, BP measures of information acquisition behavior tend to be quite consistent:
(a) across time for different subjects, different interviewers, and different (but comparable) decision tasks;
(b) within the same person, across different decision tasks; and
(c) within the same person, for the same decision task, across time (i.e., under both immediate and delayed re-test conditions). Further, BP measures seem to be at least as reliable as VR indicants. While we would agree that this assertion is based on little data, the fact of the matter is that not much data have been supplied in the consumer realm to establish the reliability of our traditional VR measures. As examples, based on a comprehensive review of more than 300 studies in the brand loyalty literature, Jacoby and Chestnut (in press) found only two studies which provided reliability data. Similarly, a fairly thorough review of the hundreds of articles comprising the advertising research literature published during the seven years spanning 1968-1974 revealed only two studies bearing on the test-retest reliability of VR measures used in that domain (Jacoby, 1976, p. 6).

Considerable additional evidence bearing on the validity and reliability of BP measures has already been collected and is being supplemented by ongoing research. These data, which include consideration of other types of validity and reliability, will be more thoroughly examined in Jacoby and Chestnut (in preparation). The primary function of this manuscript has been to provide some of the preliminary empirical evidence which suggests that BP measures demonstrate satisfactory validity and reliability and deserve a place in our methodological armamentarium. Obviously, considerably more evidence must be adduced to substantiate this assertion; we are proceeding with this task.

References

- James R. Bettman, "Information Processing Models of Consumer Behavior," Journal of Marketing Research, 7(August, 1970), 370-376.
- James R. Bettman and Jacob Jacoby, "Patterns of Processing in Consumer Information Acquisition," in B. B. Anderson (ed.), Advances in Consumer Research: Volume III. (Cincinnati: Association for Consumer Research, 1976), 315-320.
- James R. Bettman and Pradeep Kakkar, "Effects of Information Presentation Format on Consumer Information Acquisition Strategies," Journal of Consumer Research, 3 (4, 1977), 233-240.
- Donald T. Campbell and Donald W. Fiske, "Convergent and Discriminant Validation by the Multi-Trait \times Multi-Method Matrix," Psychological Bulletin, 5 (1959), 81-105.
- Robert W. Chestnut, "Information Acquisition in Life Insurance Policy Selection: Monitoring the Impact of Product Beliefs, Affect toward Agent, and External Memory," unpublished doctoral dissertation, Purdue University (1977).
- Robert W. Chestnut and Jacob Jacoby, "Consumer Information Processing: Emerging Theory and Findings," in A. Woodside, P. D. Bennett, and J. N. Sheth (eds.), Foundations of Consumer and Industrial Buying Behavior. (New York: American Elsevier, 1977), 119-133.
- Robert W. Chestnut and Jacob Jacoby, "Consumer Decision Making and Utilization of Preventative Health Care," presented at TIMS/ORSA session on Marketing Research in Health Care, San Francisco, May, 1977.
- Lee J. Cronbach, Goldine Gleser, Harinder Nanda, and Nageswari Rajaratnam, The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles (New York: Wiley, 1972). And Jack Jacoby.
- Michael J. Donahue, "Concept Formation: A Behavioral Process Approach," in preparation.
- James F. Engel, David T. Kollat, and Roger D. Blackwell, Consumer Behavior, 2nd Ed. (New York: Holt, Rinehart, and Winston, 1973).
- Jacob Jacoby, "Perspectives on a Consumer Information Processing Research Program," Communication Research, 2(3, 1975a), 203-215. Reprinted in M. Ray and S. Ward (eds.), Communicating with Consumers: The Information Processing Approach (Beverly Hills, California: Sage, 1976), 13-25.
- Jacob Jacoby, "Consumer Information Acquisition Behavior," colloquium presented to the Department of Psychology, Tilburg University, The Netherlands, May 30, 1975b.
- Jacob Jacoby, "Consumer Research: Telling It Like It Is," in B. B. Anderson (ed.), Advances in Consumer Research: Volume III. (Cincinnati: Association for Consumer Research, 1976), 1-11.
- Jacob Jacoby, "The Emerging Behavioral Process Technology in Consumer Decision Making Research," in W. D. Perrault, Jr. (ed.), Advances in Consumer Research: Volume IV. (Atlanta: Association for Consumer Research, 1977), 263-265.
- Jacob Jacoby and Robert W. Chestnut, Amount, Type, and Order of Package Information Acquisition in Purchasing Decisions, Final report to the National Science Foundation (GI-43687, June 1977).

Jacob Jacoby and Robert W. Chestnut, Brand Loyalty: Measurement and Management (New York: Wiley, in press).

Jacob Jacoby and Robert W. Chestnut, Information Acquisition in Human Problem Solving: A Behavioral Process Approach, in preparation.

Jacob Jacoby and Robert W. Chestnut, "Study of Likely Impact of Disclosure of Life Insurance Costs on Agent and Consumer Behavior," supported by Federal Trade Commission, Contract No. L0226. in progress.

Jacob Jacoby, Robert W. Chestnut, and William A. Fisher, "Simulating Nondurable Purchase: Individual Differences and Information Acquisition Behavior," Journal of Marketing Research, in press.

Jacob Jacoby, Robert W. Chestnut, and William Silberman, "Consumer Use and Comprehension of Nutrition Information," Journal of Consumer Research, 4(2, 1977), 119-128.

Jacob Jacoby, Robert W. Chestnut, Karl C. Weigl, and William A. Fisher, "Pre-Purchase Information Acquisition: Description of a Process Methodology, Research Paradigm, and Pilot Investigation," in B. B. Anderson (ed.), Advances in Consumer Research: Volume III. (Cincinnati: Association for Consumer Research, 1976), 306-314.

Jacob Jacoby, Margarete Hefner, Manfred Schöler, Klaus Grabicke, and Hans Raffée, "Information Acquisition Behavior in Brand Choice Situations: A Cross-Cultural Extension," Purdue Papers in Consumer Psychology, No. 162, 1976.

Jacob Jacoby, George J. Szybillo, and Jacqueline Busato-Schach, "Information Acquisition Behavior in Brand Choice Situations," Journal of Consumer Research, 3(4, 1977), 209-216.

Carol A. Kohn and Jacob Jacoby, "Operationally Defining the Consumer Innovator," Proceedings, 81st Annual Convention, American Psychological Association, 8(2, 1973), 837-839.

Brenda N. Major and Robert W. Chestnut, "Examining Attribution Processes with an Information Processing Methodology," in preparation.

Jum C. Nunnally, Psychometric Theory (New York: McGraw Hill, 1967).

John W. Payne, "Heuristic Search Processes in Decision Making," in B. B. Anderson (ed.), Advances in Consumer Research, Volume III. (Cincinnati: Association for Consumer Research, 1976), 321-327.

Hans Raffée, Margarete Hefner, Manfred Schöler, Klaus Grabicke, and Jacob Jacoby, "Informationsverhalten und Markenwahl," Die Unternehmung, 2(1976), 95-107.

J. Edward Russo and Larry D. Rosen, "An Eye Fixation Analysis of Multi-Alternative Choice," Memory and Cognition, 3(May, 1975), 267-276.

David A. Sheluga and Jacob Jacoby, "Do Comparative Claims Encourage Comparison Shopping? -- The Impact of Comparative Claims on Consumers' Acquisition of Product Information," in J. Leigh and C. R. Martin (eds.), Current Issues and Research in Advertising, in press.

Time magazine, March 22, 1976.

W. Fred van Raaij, Consumer Choice Behavior: An Information Processing Approach (Voorschoten, The Netherlands: VAM, 1977a).

W. Fred van Raaij, "Consumer Information Processing for Different Information Structures and Formats," in W. D. Perrault (ed.), Advances in Consumer Research: Volume IV (Atlanta: Association for Consumer Research, 1977b), 176-184.

SUBJECTS' INFORMATION PROCESSING IN
INFORMATION DISPLAY BOARD STUDIES

David C. Arch (student), University of California, Los Angeles
James R. Bettman, University of California, Los Angeles
Pradeep Kakkar, University of Pennsylvania

Abstract

The processing carried out by subjects in information display board studies is examined. A theoretical analysis of the display board task, a review of relevant prior research, and a pilot study using prompted protocol data to attempt to understand subjects' display board usage behavior are reported. The results suggest that curiosity and exploratory behavior may be widespread. Implications for further use of the display board method are considered.

Introduction

Several researchers have recently utilized information display boards to study how consumers acquire information (e.g., Jacoby, 1975, 1977; Jacoby, Chestnut, Weigl, and Fisher, 1976; Bettman and Jacoby, 1976; Payne, 1976a, b; van Raaij, 1977a, b; Bettman and Kakkar, 1977; Capon and Burke, 1977). The typical approach used in these studies has been to present information to consumers using an information display board, essentially a matrix array (with attributes as rows and brands as columns, for example). That is, each row of the board would be labelled with an attribute name for the product of interest, such as price, net weight, or unit price. Each column would be labeled with the name of a particular brand or alternative. In some studies real brand names were used (e.g., Jacoby, Chestnut, Weigl, and Fisher, 1976; Bettman and Jacoby, 1976); in others hypothetical alternatives were presented (e.g., Payne, 1976a, b). In each cell of the matrix there is a pocket or other device for holding information cards. These information cards have the appropriate brand and attribute information for that cell printed on their reverse sides. The subject is asked to choose a brand after selecting as many information cards as desired from the display board, one at a time. As each card is selected, the subject can turn the card over and read the information contained. The sequence of cards selected is then the major data of interest.

These card sequences can be analyzed in many ways. The typical approach has been to consider both how many cards were selected, and also the structure of the transitions between pairs of cards. Two general acquisition strategies used by subjects emerge from such analyses: processing by brand (examining cards for one brand, then moving on to another brand, etc.) and processing by attribute (examining cards for several brands for one attribute, then moving on to another attribute, etc.).

The display board method described above has been used to study the relationships of acquisition strategy to brand loyalty and consumption frequency (Jacoby, Chestnut, Weigl, and Fisher, 1976); effects of information presentation format on acquisition patterns (Bettman and Kakkar, 1977); effects of degree of perceived risk, amount of information, and presence of memory aids (Capon and Burke, 1977); effects of information load (Payne, 1976a, b); consumer choice heuristics (Bettman, 1977); and other phenomena related to consumer information acquisition (e.g., see Jacoby, 1977). Thus the technique has proven useful. However, the implications of some recent research lead to disquieting questions about what subjects are actually doing while acquiring information in display board studies. Accordingly, the

purpose of this paper is to examine the available evidence characterizing subjects' information processing in display board studies to raise some of the questions these research results suggest, and to assess the validity of the method in light of this evidence. Three sources of evidence are considered below: a theoretical analysis of the characteristics of the display board task; a review of prior research; and a preliminary report of results of a pilot study designed to examine subjects' processing.

A Theoretical Analysis of the Display Board Task

Newell and Simon (1972) have suggested that one can learn a great deal about the processing used by individuals in carrying out a particular task by examining the constraints imposed by that task. A consideration of what a task requires for successful completion can yield important insights into how behavior must be structured to be adaptive in that particular task environment. This theoretical consideration of task properties is termed a task analysis (see also Bettman, 1977). A brief task analysis for the information display board method is now presented.

We focus in the following and throughout the remainder of the paper on display board tasks where real alternatives are used, i.e., where actual brand names are available to the subject. This implies that various strategies are possible for the subject, including the use of memory to supplement or replace examination of the board. Different problems may arise when brand name is not available or if hypothetical alternatives are used (e.g., prior knowledge may not be as great an influence on processing), but consideration of such problems is beyond the scope of this paper. Several aspects characterizing the display board task with known brand names are now considered: the interaction with memory; format considerations; how information is accessed; the types of acquisition possible; the effort required for acquisition; and the obtrusiveness of the method.

Subjects bring some prior knowledge to the task if known brands are used on the display board. This prior knowledge may interact with acquisition. The subject may acquire less from the board and use more information from memory; the subject may check up on information in memory by acquiring the relevant information from the board; the subject may acquire information from the board to see if it looks plausible given what is in memory (i.e., may check up on the experiment); and so on. Thus, memory can impact not only how much information is acquired, but also the purposes underlying some observed acquisition sequence. This interaction of the information acquired with that in memory needs to be examined more closely.

The matrix format used in display board tasks allows either brand or attribute processing to be carried out easily. However, almost all real world information acquisition settings place considerably more constraints on processing. Shelf displays, package information, and many ads are organized by brand. That is, information

is presented separately, for one brand at a time. Thus brand processing would be easier than attribute processing. In addition, display boards present the subject with a very well-structured, comprehensive, and organized information display. Although some real world displays are also well-organized (e.g., a table in Consumer Reports), in most cases consumers must piece together information from several different sources, presented in different fashions, perhaps on different attributes for different brands (e.g., acquiring information from television commercials or from manufacturers' brochures). The implication is, therefore, that the format and organization of the display board task does not place the same constraints on acquisition as do most real world environments. For more discussion see Bettman (1977).

The accessing of information also appears to differ between display board tasks and actual consumer information displays. In the display board task, subjects can only acquire one piece of information at a time: acquisition is sequential. In looking at a package or some other displays, however, consumers can see several pieces of information at the same time: acquisition can be (essentially) simultaneous. For example, in examining a nutritional label on a cereal package, a consumer can see the general pattern of the nutrient percentages simultaneously.

Display board tasks tend to focus on intentional or goal-directed information acquisition. That is, for a consumer to acquire a piece of information in a display board study, he or she must specifically pull out a card from the board. This may be contrasted with a consumer's scanning a package visually and "happening" to see some piece of information. This more serendipitous type of acquisition is not possible with display boards, but may be very characteristic of actual information acquisition.

To examine another related property of display board tasks, let us assume that the display board and some visual information display (e.g., a package label or a table) contain the same information. Then the effort required to obtain information from the display board may be relatively high compared to visual search of the other information display. That is, in the display board task, the subject must remove the various cards desired, rather than simply scanning the display. This analysis would also imply that information acquisition would potentially be much slower for display board tasks than for visual search of the same information displayed in a table or on a label. Thus less information in total might be selected in display board tasks than in cases where a package or other display were visually examined. Use of information in memory might also be more widespread in display board tasks than in visual search, due to the greater effort involved in directly examining information from the board (Russo, 1978a), as compared to the effort involved in visual search.

Several of the aspects discussed above can now be brought together to analyze the relative difficulty of acquiring information from a display board or from some real world display. The matrix format and well-structured, comprehensive nature of the display board would tend to make acquisition easier from the board. The greater effort and time required for acquisition from the board and the limitation to sequential rather than simultaneous access would tend to make acquisition from the board harder. Therefore, the difficulty of the display board as compared to some real world display would depend upon which of the above factors is stronger. Some actual information displays would be harder than the board (e.g., cases where information was not available in one place, but had to be pulled together by the consumer, where the information was presented in different formats for different brands, and so on). Thus format and the availability of information in a comprehensive

and well-organized display might dominate the effects of how information could be accessed. On the other hand, some actual displays would be easier than the board (e.g., if the same information available on the board were displayed in a Consumer Reports table or on a package label). In this case, the relative ease of visual search compared to choosing cards would dominate the format factor, since the format and organization would be essentially the same for both displays. Even in these cases, however, simply comparing the relative difficulty of the display board and some real world acquisition setting can be misleading, because the factors underlying the difficulties encountered in each task may be very different, as the above analyses demonstrate.

Finally, display board methods are very obtrusive. Subjects may think that the information they acquire is being carefully monitored, and therefore may try to select more "rational" pieces of information. The analyses above suggest that there may be some biases in the display board task. We now turn to an examination of selected prior research using display boards to determine if these studies can shed further light on subjects' processing during such tasks.

Prior Research on Subjects' Processing in Display Board Tasks

Three areas of research which have implications for the type of processing carried out by subjects in display board tasks are considered below: effects of information presentation format; comparison of display board and eye movement acquisition results; and effects of time pressure.

Information Presentation Format Effects

As noted above, display board tasks, with their organized matrix format, may make some kinds of processing (e.g., attribute processing) easier than in actual choice environments. Bettman and Kakkar (1977) tried to directly study the impact of format on acquisition. Subjects were 150 housewives, divided equally into three groups. One group (Matrix) received information on cereals in the standard matrix display board format. A second group (Brand) received information in separate booklets for each brand, and a third group (Attribute) received separate booklets for each attribute. Thus the Brand group's format should encourage brand and discourage attribute processing, with the opposite true for the Attribute group. The results were strongly supportive of the predictions. Subjects in the Brand group used almost exclusively brand processing, and subjects in the Attribute group used attribute processing extensively. Subjects in the Matrix group used both brand and attribute processing.

These results suggest that major variations in display format can have a strong impact on acquisition. They also suggest that the matrix format used in display boards may allow or encourage different acquisition strategies than those characteristic of some real world environments. Thus the theoretical format bias discussed above has empirical support.

Comparing Display Board and Visual Package Search

Several of the theoretical analyses above implied that there might be important differences between acquisition from a display board and visual search of a package or some other information display. It was suggested that if the same information were available from the board and from a visual display, then more information might be acquired and acquisition might be faster for visual search. Van Raaij (1977a, b) reports a study which directly compared display board and visual search

behavior.

In Van Raaij's study (1977b), the same twenty housewives participated in an eye movement study and then in a display board study four months later. The same thirteen alternative brands of coffee were used in both studies, each characterized by four attributes. In the eye movement study, actual product packages with the four attribute values affixed to the sides were examined by the subjects, and their eye movements were recorded. The display board study used a standard matrix display. Van Raaij examined both brand name present and brand name absent conditions; only the brand name present study is considered below.

Several suggestive results were obtained. First, subjects acquired more information in the eye movement study (26.2 versus 15.7 different items -- these figures do not count examining the same piece of information more than once as a separate acquisition). Second, of the pieces of information they selected, subjects examined 51.4% twice or more in the eye movement study, but there was no reacquisition in the display board study. These two findings provide support for the notion that acquisition during visual search may in many cases be more accidental, a result of scanning the package, while acquisition from a display board tends to be intentional or consciously goal-directed.

A third finding was that the acquisition sequence measures were very similar for both studies, with more brand processing evident. Finally, the choice times were greater in the display board conditions than in visual search. Hence, more information was examined per unit time in visual search than in the display board study (Van Raaij, 1977a). The results described above imply that there may be some major differences between visual package search and acquisition from a display board in situations where the same information is available on both the display board and the package.

Effects of Time Pressure

Bettman and Kakkar, in an unpublished study, examined the effects of time pressure on acquisition in a display board task. Subjects were 150 housewives, equally divided into three conditions, all using the same display board task for breakfast cereals. In the first condition (No Pressure), subjects were given no instructions about the time they might take to complete the task. In a second condition (Full Time Asked), subjects were told that, on average, other subjects took two minutes to complete the task. They were then asked how much time they would need, and were given the full amount asked for (the average time asked for was 114 seconds). In the third condition (Half Time Asked), subjects were also told that other subjects took two minutes and were asked how much time they would need (the average time asked for was also 114 seconds). However, this group was then told that due to the number of subjects to be run, only half the time they asked for could be allowed.

The results were somewhat surprising. As expected, the time taken to choose a brand differed across the three groups ($p < .001$). The No Pressure group took 108.5 seconds, the Full Time Asked group 69.7, and the Half Time Asked group 45.3. However, there was not a significant difference in cards taken ($p = .147$) although there was a trend in the expected direction (No Pressure - 7.4; Full Time Asked - 6.5; Half Time Asked - 5.1). There were significant differences in time taken per card ($p < .01$), with the Half Time Asked group taking the least time and the No Pressure group the most. One other unexpected result was that although the Half Time Asked group felt more hurried than the No Pressure group, the Half Time Asked and Full Time Asked groups

did not differ in feelings of being hurried or being under time pressure. Finally, the acquisition sequence measures (types of transitions) did not differ across groups.

These results seem to imply that subjects adapted to the attempt to impose time pressure by simply speeding up their processing. This may mean that subjects could speed up because they were checking up on things they already knew. If this were the case, it might raise serious questions about subjects' goals for examining information in display board studies. However, it might also mean that subjects were processing in a fairly leisurely fashion in the No Pressure and Full Time Asked groups, and that there was a good deal of "slack" time available. Hence, subjects might have been able to speed their processing easily, without really burdening themselves. This speeding up observed could be a method generally used to adapt to tasks when time available for processing is decreased. If time pressure were made quite severe, subjects might not be able to adapt by speeding up, and might have to resort to other strategies (e.g., restricting the brands and attributes they examine). However, Bettman and Kakkar apparently did not impose this much time pressure with their manipulations.

The theoretical analyses and review of prior research have indicated that there may be some serious biases or problem areas in display board tasks. The major issues appear to be: (1) The impact of prior knowledge about the alternatives on acquisition from the board; (2) Impacts of format on acquisition, with some evidence that display board formats may encourage or allow different kinds of processing than some kinds of real world environments; (3) Apparently substantial differences between display board and visual search acquisition, including how much information can be simultaneously accessed, whether acquisition can be accidental or is more goal-directed, how much time is taken, and how often information is re-examined; and (4) Impacts of time pressure on acquisition.

The above discussion has presented only an aggregate view of subjects' processing in display board tasks, and the picture is complex. It is still very difficult to characterize what subjects are actually doing while they are acquiring information from a display board. In particular, there is very little direct knowledge of the goals underlying subjects' information acquisition or of the effect of subjects' prior knowledge on acquisition. Thus there is limited hard evidence about details of subjects' processing, although there are many possible hypotheses (not necessarily mutually exclusive). For example, subjects may be (1) Acquiring information as they would in a store; (2) Checking to see if the information in the display board is the same as they think it should be (i.e., seeing if the experimenter is "tricking" them); (3) Exploring brands they know very little about, or exploring brands used previously but not currently; (4) Using the opportunity to "educate" themselves about their current brand; and so on.

An exploratory study was carried out by Arch and Bettman to attempt to gain some more detailed insights into subjects' processing. Preliminary results of that study are now reported.

An Exploratory Study of Subjects' Processing in Display Board Tasks

The study used two main approaches to try to increase understanding of subjects' processing. First, an attempt was made to measure prior knowledge for the brands used. Second, prompted protocols were obtained from subjects after they had selected their information cards.

Method

Subjects were ten female shoppers, ranging in age from 18 to 49. They were recruited by interviewers in a shopping mall and asked to answer some questions about consumer products. They were told this would take about 20 minutes, and were paid for their time.

Subjects were led by an interviewer into a partitioned area within the mall, where they answered questions about frequency of purchase and familiarity with breakfast cereals. Subjects were then asked to fill out a Brand Characteristic Summary, in an attempt to gather fairly structured information about prior knowledge. The Brand Characteristic Summary was a table with eleven characteristics (attributes) for cereal listed along the rows, and ten brands of cereal listed along the columns (thus there were 110 cells or boxes in the table). The subject was instructed to put a 1 in those boxes corresponding to characteristics of brands that they knew a lot about and would not need to check in the store; a 2 if they knew a little or a fair amount but would want to check it in the store; a 3 if they knew almost nothing but would want to check; and to leave the boxes blank for those brands and characteristics they would not want to check even if they knew little. This task was designed to elicit what pieces of information subjects might check while in the store shopping, so that the information acquired later in a display board task could be compared to it. There may be some bias in the Brand Characteristic Summary task, in that subjects may overstate what they know and what they claim they would want to check in the store. This is discussed more fully below.

Since subjects were later to be given a display board task for cereals, some intervening tasks were included between the Brand Characteristic Summary and the display board to diminish possible attempts at consistency by subjects trying to appear "rational." Accordingly, subjects also filled out some questions and a Brand Characteristic Summary for steam irons. Then the subjects went to another table where they tried samples of lip balms and answered several questions about these samples.

Subjects were then given a standard display board task for cereals, using the same attributes and brands as in the Brand Characteristic Summary. No subject connected these two tasks in post-task debriefings. After completing the display board task and choosing a brand, subjects answered several questions about their reactions to the task (e.g., certainty they had made the best choice, confusion with the task, whether they felt they behaved the same as in the store, which brands of cereal they purchased most frequently). While the subject was doing this, the interviewer took the sequence of cards selected by the subject from the display board, and put X's in the boxes of a new Brand Characteristic Summary table for those cards selected.

After the subject finished the questionnaire, the interviewer gave her the deck of cards she had selected and the table marking those cards, and told her that the X's on the table represented the cards chosen. The interviewer then asked the subject to go through the deck of cards in the order they were originally selected and to talk about what was going through her mind when she first took these cards. This verbal report, which was tape recorded, has been called a prompted protocol (e.g., Russo and Rosen, 1975), prompted by the sequence of cards taken. Such prompted protocols are related to, but not the same as, protocols taken simultaneously, during the actual acquisition task itself (e.g., Payne, 1976a, b). Prompted protocols differ from simultaneous protocols in several ways. First, since the prompted protocols are taken after the display board task has

been completed, there are no problems with the protocol's interfering with or biasing the acquisition behavior observed. Second, the prompted protocols have proven useful in other studies (Russo and Rosen, 1975) at providing a somewhat more structured view of the processing undertaken than is usually obtained with simultaneous protocols. However, there is also potentially a major disadvantage to prompted protocols, in that, since they are not taken while the behavior is occurring, there is more danger of rationalization or retrospection about "why I did that" (Nisbett and Wilson, 1977). On balance, it was felt that the somewhat more structured view provided by prompted protocols and the fact that the acquisition behavior would not be biased were more overriding concerns in the present study.

Following the prompted protocols, the interviewer took out the original Brand Characteristic Summary and compared that to the cards taken, asking the subject about selected differences (e.g., items marked with a 1 on the Brand Characteristic Summary (know a lot, don't need to check) but then acquired, items marked with a 2 or 3 (know some or a little, need to check) but later not acquired, and so forth). This was also tape recorded.

The description of the procedure is lengthy. To summarize, the major data available are measures of purchase frequency and familiarity, Brand Characteristic Summary information, the sequence of cards selected from the display board, reactions to the display board task, prompted protocols taken after the display board task, and some probes for differences between the Brand Characteristic Summary and display board tasks. One can then examine characteristics of the display board acquisition patterns, characteristics of the Brand Characteristic Summary data, the relationship between the Brand Characteristic Summary measures and the information acquired from the display board, and the prompted protocol data.

Results

Display Board Results. In the aggregate, the display board results agree with prior research. Subjects selected an average of 10.2 cards, or roughly 9% of the 110 available. Subjects' acquisition sequences were classified, using the criteria in Bettman and Kakkar (1977). Seven subjects used brand processing, two used attribute, and one used another strategy. This agrees with the Bettman and Kakkar results, which showed that more experienced shoppers tend to use more brand processing. Subjects also seemed reasonably certain they had made the best choice (a mean of 5.2 on a 7-point certainty scale) and claimed that their behavior on the display board task was very much like their in-store behavior (a mean of 6.2 on a 7-point scale with endpoints 1 = Not at all like I behave in the store and 7 = Exactly like I behave in the store). This latter result should certainly be viewed with some scepticism, since subjects' responses may be biased upward by desires to please the experimenter, i.e., to let the experimenter know they were cooperative.

Brand Characteristic Summary Results. Subjects claimed to have a good deal of prior knowledge in the Brand Characteristic Summary. On average, subjects checked 18.1 items with a 1 (know a lot, wouldn't need to check); 22.4 items with a 2 (know some, would want to check); and 28.8 items with a 3 (know little, would want to check). Thus subjects marked an average of 69.3 items of 110, or 63%. Also, subjects claimed they would check in the store 51.2 (the sum of the 2 and 3 responses) of 110 items, or 46.5%.

Comparing the Display Board and Brand Characteristic Summary Results. At the aggregate level, it is obvious that subjects claimed in the Brand Characteristic

Summary that they would check a great deal more information than they actually did examine in the display board task -- 51.2 items (46.5%) in the Brand Characteristic Summary versus 10.2 items (9%) in the display board task. In addition, not all of the cards selected in the display board task correspond to responses of 2 or 3 (know some or a little, want to check) in the Brand Characteristic Summary. In fact, of the average of 10.2 cards taken from the display board, 3.6 were items checked with a 1 (know a lot, don't need to check) in the Board Characteristic Summary; 2.8 were items checked with a 2 (know some, want to check); 2.6 were checked with a 3 (know little, want to check); and 1.2 were blank (don't want to check). Thus subjects actually selected cards for a higher percentage of the items they had marked with a 1 (and hence said they wouldn't need to check) than for items they had marked with a 2 or 3 (and hence said they would want to check).

These results show that subjects sometimes do not acquire information in the display board task, even though they claimed they would in the Brand Characteristic Summary. Subjects actually acquired at least one card in the display board task for 51% of the attributes and for 44% of the brands they checked in the Brand Characteristic Summary.

The results also indicate that subjects acquire information they earlier claimed they would not need to examine. Those cases where subjects marked items with a 1 in the Brand Characteristic Summary and then selected the corresponding cards in the display board task were examined, using subjects' tape recorded responses to the interviewers' probes. Subjects' stated reasons varied. Some of these instances simply appear due to unexplained inconsistency in the subjects' responses; some appear to be due to subjects' checking items to see if the information is what they believe it to be; and some are due to subjects' checking attributes that they feel they know for their store, but which might vary from store to store (e.g., price). One other reason appears which is quite interesting, and which may shed some light on the differences between the Brand Characteristic Summary and display board results. Subjects may say they know a lot and don't need to check when they know a characteristic's general level -- e.g., they know protein content is 'high.' They may then later check that item to refine their knowledge, to find out that 'high' protein is 10 grams. Thus 'knowing' a lot may mean knowing general ranges to the subject, not knowing an exact amount.

The discussion above amply documents that there are substantial disagreements between subjects' claims in the Brand Characteristic Summary and their acquisition behavior in the display board study. There are several possible reasons for this. First, it may be that the Brand Characteristic Summary task encourages overstatement of intended search, since this appears rational. Second, the Brand Characteristic Summary, as asked, may not have been seen by the subjects as being specific to their own shopping behavior. That is, rather than responding with what they would actually want to check if shopping for cereal in the store, subjects may have responded with what they might check if they were to consider some brands in the future, or what they might check if they had time. On the other hand, subjects made many comments in the protocols characterizing the display board task as a shopping task (which of course does not necessarily mean it was really treated that way). Although the evidence is scanty (mostly subjects' comments), the second reason seems more likely.

Insights from the Prompted Protocol Data. The contents of the prompted protocol data were judged by the authors. In our future research on this topic, with more subjects, independent judges will be used. Thus these

preliminary results are only suggestive. The prompted protocol data suggest two main findings. First, there seems to be a great deal of exploratory or curiosity-based information acquisition in the display board task. Subjects spent a good deal of time checking on brands they had previously purchased but no longer bought, saying they did this because they were curious about what these old brands were really like (this occurred for five of ten subjects). Subjects also spent time exploring new brands they knew little about (for four of ten subjects). One of these subjects said, "I didn't know too much about the cereals I did check. That is why I was going through them..." Finally, subjects also used the display board task to explore or educate themselves about their current favorite brand (for five of ten subjects). Thus there is curiosity about brands previously bought, relatively unknown brands, and the consumer's current brand. In many of the cases where exploratory behavior was observed, there appeared to be little explicit comparison of alternatives carried out.

A second finding is that some display board acquisition appears to be aimed at checking on the experiment (this may have occurred for three of the ten subjects). As one subject stated, "I wondered what your card was going to say."

Discussion

The results of the exploratory study suggest two broad areas for discussion: how to measure prior knowledge, and how to interpret the curiosity findings.

The discrepancies noted above between the display board and Brand Characteristic Summary tasks suggest that more work needs to be done on the measure of prior knowledge used. One must first distinguish between general knowledge (protein is high) and specific knowledge (protein is 10 grams). Subjects appear to interpret the Brand Characteristic Summary as asking for general knowledge. The Brand Characteristic Summary may have failed as a good measure of this general knowledge because subjects did not relate it to their actual shopping behavior, but responded more with what they might do sometime. Thus, one might strengthen the Brand Characteristic Summary by being very specific, attempting to tie it to shopping behavior (e.g., asking subjects to indicate what they would do in the store the next time they shopped for cereal, or what they did the last time). One might try to validate such a measure (and the display board findings as well) by measuring actual search behavior, perhaps in an artificial store setting, but this may be very difficult. For specific, as opposed to general, knowledge, one may have to test actual values. It is not clear whether measures of general or specific knowledge would be more related to acquisition behavior.

The curiosity findings in the display board task are intriguing. They may, of course, be spurious. One could argue that this exploratory behavior is a result of low involvement with the task. Subjects may decide that they are in an experiment, they might as well get something out of it, and that the situation is one which is well-suited for learning about new brands, old brands, or one's current brand, since an array of information is available. On the other hand, these findings could be real. There is very little evidence about exploratory behavior in actual choice situations, a major gap in consumer research. However, subjects claim that they treated the board like a shopping situation. Although this is weak evidence at best, the exploratory behavior results seem definitely worth pursuing.

Overall Discussion

Three sources of evidence about display board tasks have been considered: theoretical arguments, prior research,

and the preliminary study reported above. The evidence itself is very hard to characterize; however, some conclusions seem warranted. First, display board tasks appear to have some serious biases, which may limit the range of phenomena they can be reasonably used to study, but also possess some strengths for researching other areas. For example, display board tasks may be very useful for researching curiosity-based or exploratory behavior, a phenomenon which has proven extremely difficult to research. Also, display board tasks may be very suited for studying the goals consumers develop for information acquisition. That is, the consumer may set up some goals for search, for what information to examine. The information acquired in a display board task may relate to these goals, since intentional or goal-directed search is largely what is observed.

Display board tasks, because of their biases, do not seem as suitable for some real world visual search tasks. As noted in the Van Raaij (1977a, b) findings above and in the theoretical task analysis, the display board task diverges from some visual search tasks in many crucial respects. In particular, the above arguments imply that display board tasks may not be appropriate for studying acquisition of package information, especially where fairly complete information is available on the package. As a specific example, results of display board studies might be particularly misleading for drawing public policy implications about providing package information. One major aim of providing information on a package may be to have consumers learn about that information over time by "accidentally" acquiring it while scanning the package. As noted earlier, display board tasks do not allow for this type of phenomenon to any great extent. Thus actual acquisition of information over time and its eventual impact may be severely understated by results of display board studies. Eye movement analyses seem best suited to studying most visual search tasks (Russo, 1978a, b).

One final comment seems necessary. Both the display board task and eye movement studies suggested above are obtrusive tasks. There is probably a consumer research counterpart to the Heisenberg Uncertainty Principle: one cannot measure process without impacting that process. Thus process methodologies will have impacts; the problem is to understand those impacts.

References

James R. Bettman, "Data Collection and Analysis Approaches for Studying Consumer Information Processing," in William D. Perreault, ed., Advances in Consumer Research, Volume IV (Atlanta: Association for Consumer Research, 1977, 342-48).

James R. Bettman and Jacob Jacoby, "Patterns of Processing in Consumer Information Acquisition," in Beverlee B. Anderson, ed., Advances in Consumer Research, Volume III (Cincinnati: Association for Consumer Research, 1976, 315-20).

James R. Bettman and Pradeep Kakkar, "Effects of Information Presentation Format on Consumer Information Acquisition Strategies," Journal of Consumer Research, 3 (March 1977), 233-40.

Noel Capon and Marian Burke, "Information Seeking Behavior in Consumer Durable Purchase," in Barnett A. Greenberg and Danny M. Bellenger, eds., Contemporary Marketing Thought (Chicago: American Marketing Association, 1977, 110-15).

Jacob Jacoby, "Perspectives on a Consumer Information Processing Research Program," Communication Research, 2

(July 1975), 203-15.

Jacob Jacoby, "The Emerging Behavioral Process Technology in Consumer Decision-Making Research," in William D. Perreault, ed., Advances in Consumer Research, Volume IV (Atlanta: Association for Consumer Research, 1977, 263-65).

Jacob Jacoby, Robert W. Chestnut, Karl C. Weigl, and William Fisher, "Pre-purchase Information Acquisition: Description of a Process Methodology, Research Paradigm, and Pilot Investigation," in Beverlee B. Anderson, ed., Advances in Consumer Research, Volume III (Cincinnati: Association for Consumer Research, 1976, 306-14).

Allen Newell and Herbert A. Simon, Human Problem Solving (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1972).

Richard E. Nisbett and Timothy DeCamp Wilson, "Telling More Than We Can Know: Verbal Reports on Mental Processes," Psychological Review, 84 (May 1977), 231-59.

John W. Payne, "Heuristic Search Processes in Decision Making," in Beverlee B. Anderson, ed., Advances in Consumer Research, Volume III (Cincinnati: Association for Consumer Research, 1976, 321-27).

John W. Payne, "Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis," Organizational Behavior and Human Performance, 16 (August 1976b), 366-87.

J. Edward Russo, "Adaptation of Cognitive Processes to the Eye Movement System," in J. W. Senders, D. F. Fisher, and R. A. Monty, eds., Eye Movements and the Higher Psychological Functions (Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1978a).

J. Edward Russo, "Eye Movements Can Save the World: A Critical Evaluation," in H. Keith Hunt, ed., Advances in Consumer Research, Volume V (Chicago: Association for Consumer Research, 1978b).

J. Edward Russo and Larry D. Rosen, "An Eye Fixation Analysis of Multi-Alternative Choice," Memory and Cognition, 3 (May 1975), 267-76.

W. Fred van Raaij, Consumer Choice Behavior: An Information Processing Approach (Voorschoten, The Netherlands: VAM, 1977a).

W. Fred van Raaij, "Consumer Information Processing for Different Information Structures and Formats," in William D. Perreault, ed., Advances in Consumer Research, Volume IV (Atlanta: Association for Consumer Research, 1977b, 176-84).

EYE FIXATIONS CAN SAVE THE WORLD: A CRITICAL EVALUATION AND A COMPARISON
BETWEEN EYE FIXATIONS AND OTHER INFORMATION PROCESSING METHODOLOGIES

J. Edward Russo, University of Chicago

Abstract

Eye fixations are described and evaluated as a type of process tracing data. Several examples illustrate their use, in both laboratory and consumer settings. This method is compared with four alternatives: chronometric analysis, information display boards, input-output analysis, and verbal protocols. Each method is evaluated on seven performance attributes ranging from validity to equipment cost. The resulting ratings are summarized in a single table. Each method is seen to have unique advantages, so that no one method dominates and no one method should be shunned. The use of joint methodologies is strongly advocated. Eye fixations and verbal protocols are especially complementary methods.

Introduction

The phenomenological base of the information processing paradigm is the sequence of cognitive events that occur between stimulus and response. The ascension of this paradigm in consumer research has shifted the theoretical concern from states to processes and the empirical focus from the output of a process to data that can identify the process itself. To identify the sequence of events that constitute information processing, new methodologies have been developed and old ones elaborated. The goal of this paper is to describe one of these methodologies, the recording of eye fixations, and to compare it to the most popular alternative techniques.

Process Tracing Data

The available methodologies can be classified as process tracing or otherwise. All process tracing techniques rely on intermediate responses to provide evidence of ongoing cognitive activity. These responses occur during the performance of the task, prior to the availability of the task's output. Process tracing methodologies include the recording of eye fixations, verbal protocols, and information display board techniques. Input-output analysis and chronometric analysis are examples of methodologies that do not rely on a trace of the process.

The intermediate responses that constitute the process trace can be either naturally occurring behaviors, like eye fixations, or they can result from additional instructions, like verbal protocols. There are many possibilities in both categories. For example, natural overt behaviors include the use of external memory, such as writing intermediate computations for later retrieval. Instructed behaviors include, besides verbal protocols, the category of subgoal markers. For example, when choosing from several brands, consumers can be required to mark each brand as it is eliminated. These overt eliminations provide a trace of the progress of the choice task. The potential range of process tracing observations is very large. Only a few of these possible data collection techniques are now in use in any of the behavioral sciences.

Eye Fixations

For humans the most efficient means of obtaining information from the environment is through eye fixations.

Fixation points are located very accurately, and the intervening movements of the eyes are very fast. For example, a typical 5-6 deg. movement requires about 30 msec. This means that during typical tasks, only 5% of the time is devoted to eye movements, while useful fixation time accounts for the other 95%. The speed and accuracy of the visual system are unrivaled by other human sensory systems.

Eye Fixations as Information Acquisition Responses

When using eye fixations to trace cognitive processes, one limitation must be kept in mind. Eye fixations are information acquisition responses. Rarely are they anything else. This means that using them to identify cognitive processes is like using only the tip to judge the shape of an entire iceberg.

Cognitive processing can be partitioned into two types of activity, the acquisition of information and the operations that are performed on this input. Any strategy for performing a cognitive task, such as consumer decision making, will exhibit a characteristic pattern of information acquisition and internal computation. The task of the researcher is to identify the consumer's strategy from only what is observable. The internal computation is not observable, and neither is any information that is retrieved from memory. It is mainly when the guiding strategy calls for external information that observable behavior is produced.

For a researcher to infer the underlying cognitive strategy from a sequence of eye fixations, the task environment must be structured around the acquisition of visual information. Even then this inference is seldom an easy task, and it is sometimes impossible. However, cognitive psychologists and other behavioral scientists have had some notable recent successes in eye fixation analysis (e.g., Just and Carpenter, 1976a). These data are becoming more useful as we understand better how to interpret appropriate subsequences of fixations and as process-oriented theories become available to guide the examination of these data.

Collecting Eye Fixation Data

There is no single best way to collect eye fixation data. Many different techniques are available. These techniques differ widely in convenience, price, subject restraint, and other relevant attributes. Each of the recording methods has characteristic advantages that can make it well suited to some experimental situations but not to others. For example, to gather pilot data, direct observation of the subjects' eyes may provide enough accuracy at a low cost. For a confirmatory study involving detailed stimuli like package facings, a more precise, computer-controlled eye position sensor may be needed.

Simple Observation. Direct observation methods are the least complicated way to collect eye fixations. Typically, a video recording of the subject's face is made during task performance. At some later time, this recording is coded into a sequence of eye fixations by combining knowledge of the stimulus display with a subjective determination of the subject's direction of gaze. The latter task can be time consuming, especially if many subjects are involved. It is also susceptible

to a trade-off between reliability and accuracy. The primary advantages of direct observation techniques are minimal set-up and equipment costs. (The only equipment needed is a video tape recorder and playback monitor.) Furthermore, using this method, eye fixations can sometimes be recorded unobtrusively. For example, a one-way mirror can be used to separate the video camera from the subject (Van Raaij, 1976a; Russo, 1977a).

Precise Recording. If large amounts of precise eye fixation data are needed, a computer-controlled system is essential. Such a system consists of two major components: an eye position sensor and a laboratory computer for recording and analyzing the fixations. These systems can be quite expensive, but they offer great accuracy and efficiency. With appropriate software, a print-out of the location, duration, and sequence of eye fixations is available immediately (Loftus, Mathews, Bell, and Poltrock, 1974). Furthermore, on-line analysis of eye fixations offers exciting possibilities for altering the visual display as a function of where the subject is looking. For examples of such systems and their use, see Just and Carpenter (1976b) and Reder (1973).

Further Reading

The preceding brief overview of eye fixation methodology is necessarily selective. The following readings are recommended to readers seeking more detail on the recording and analysis of eye fixations.

A forthcoming volume by Senders, Fisher, and Monty (1978) contains an up-to-date survey of the use of eye fixations in cognitive psychology and several applied disciplines. In spite of an emphasis on military applications, this book is highly recommended. Included in this compilation is a paper by the present author that analyzes in detail the relation between eye fixations and underlying cognitive strategies (Russo, 1978). Also recommended is a similar, but earlier, volume by Monty and Senders (1976). The classic reference work on the eye is a six-volume collection of papers edited by Davson (1962-1975). Particularly relevant is the chapter on eye movements by Alpern (1969).

For a review of data collection methods and apparatus, a recent comprehensive survey by Young and Sheena (1975a, b) is available. The simpler, less expensive methods are reviewed by Russo and Mathews (1975). Just and Carpenter (1976b) describe a complete eye movement laboratory and its use. Examples of software systems that are used in computer-controlled eye fixation recordings are provided by Conery, Smith, and Russo (1974), Reder (1973), and Goode and Russo (1971).

Examples of the Use of Eye Fixations

The best way to explain the methodology of eye fixations is by example. Two complementary applications have been selected from the author's personal experience (Russo and Rosen, 1975; Russo, 1977a). Both focus on the multi-alternative choice process, first in a setting characteristic of experimental psychology and then in a consumer setting. When examining these examples, the reader is urged to consider the general problems of interpreting eye fixation sequences, rather than the specific interpretations presented. These general problems include validating an interpretation and isolating the relevant cognitive units of behavior.

Multialternative Choice

Eye fixations were used by Russo and Rosen (1975) to expose the strategies used in multialternative choice. The subjects were college students and the choice sets were always six used cars. The cars were described by

three alphanumeric attributes, make, year, and mileage; no pictures were used. A preliminary measurement of utility permitted the construction of choice sets of uniform difficulty both within and across subjects. The eye fixation sequence from a typical trial is presented in Table 1. These data include refixations only; the initial fixation on each alternative was controlled by the stimulus display.

TABLE 1
REPRESENTATIVE HYPOTHETICAL SEQUENCE OF
EYE FIXATIONS FROM A MULTIALTERNATIVE
CHOICE TASK

Refixation Number	Alternative Number
1	3
2	2
3	1
4	4
5	1
6	5
7	2
8	3
9	6
10	3
11	6
12	3
13	6
14	4
15	6
16	4
17	1
18	2
19	6
20	3
21	6
22	3
23	2
24	4
25	3

In analyzing this sequence of fixations, the goal is to link these information acquisition responses to an underlying cognitive strategy. The aspect of these data that first strikes an observer is the pattern of alternation between two alternatives. There are four such alternating chains of fixations, marked in Table 1 by vertical brackets. They range from a minimum of three to six fixations in length. This pattern of eye fixations suggests a series of pair comparisons. That is, it suggests that the subject's strategy for multialternative choice is to partition the table into a sequence of pair comparisons. Over twelve subjects, about 64% of refixations were in pair comparisons, with only one subject exhibiting less than 50%.

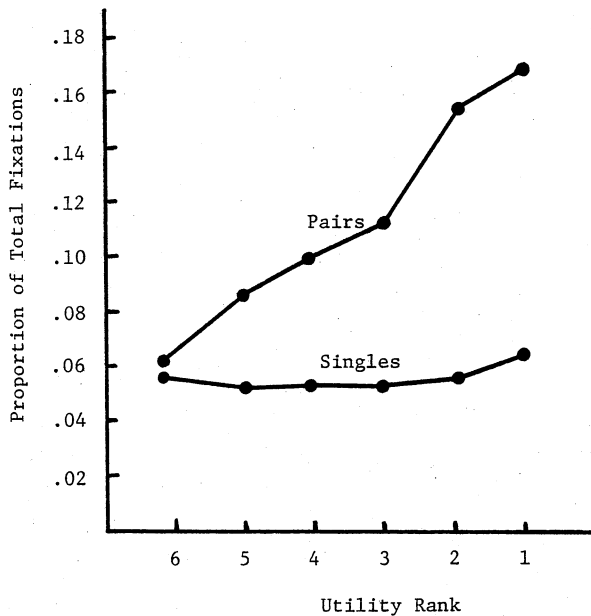
The interpretation of an alternating sequence of three or more fixations as a pair comparison may have face validity. But how can the researcher be certain that this interpretation is correct? Two validation procedures were used. First, verbal protocols were collected, and these data were examined for statements that confirmed the interpretation of eye fixation sequences. When the alternating pattern included four or more fixations, a 95% confirmation rate was obtained. That is, for 95% of the longer alternations, the subjects explicitly stated that the two fixated alternatives were being compared. When there were only three fixations, an X-Y-X pattern, the confirmation rate dropped to 69%. This suggested that some of the briefer alternations did

not indicate the occurrence of pair comparisons. On average, 83% of the fixation alternations were accompanied by explicit statements of comparison. Since the verbal protocols could not be expected to be complete, this confirmation rate is a lower bound.

A second validation procedure relied on the prior measurement of the utility of each of the six alternatives. A strategy of pair comparisons implied that evaluation took place within these alternating patterns of eye fixations. However, evaluation might also have taken place during the remaining "single" fixations. Figure 1 implies that this was not the case. All evaluative processing took place in the pair comparisons.

FIGURE 1

RELATIVE FREQUENCY OF REFIXATIONS BY UTILITY RANK



To understand Figure 1 consider the following: whenever evaluation is taking place there should be a sensitivity to the utility of the alternatives. The higher the utility of an alternative the more attention (number of fixations) should be received. This predicted pattern holds only for the pairs. The singles show no sensitivity to utility, indicating that no evaluative processing took place during these fixations.

Both procedures for validating the interpretation of the eye fixations required the use of additional data. First, verbal protocols directly confirmed the hypothesized interpretation. Second, a series of outputs of the evaluative process were transformed into utility ranks, which had a uniquely predicted relation with the eye fixations. Validation procedures are almost always examples of the convergent use of two methodologies for examining the same phenomenon.

Behavioral Units versus Cognitive Units

This example illustrates a critical step in the analysis of eye fixations. Subsequences of eye fixations, namely alternations of three or more fixations, were interpreted as single cognitive units. Although the behavioral unit

was a single eye fixation, the cognitively meaningful unit was a group of eye fixations. It is essential to aggregate the behavioral units into the appropriate cognitive unit (see Just and Carpenter, 1977a). Consider the interpretation of fixation durations. In this experiment great effort was expended trying to interpret the durations of individual fixations, especially within a single pair comparison. No success was achieved, however, because these single fixations did not represent complete cognitive units. In contrast, the durations of the pair comparisons yielded important insight into the subjects' strategies.

Another way of emphasizing the importance of identifying cognitive units is to consider what happens when they cannot be found. The early literature on reading contains many examples of extensive recording of eye fixations, often with primitive equipment and great labor cost. For an example, see Kolers (1976). These studies relied on summary statistics, such as the location, frequency, and mean duration of the observed eye fixations. Although such measures are informative about reading performance (e.g., good readers need fewer fixations), they say little about the reading process, i.e., the process of comprehension. As in many other situations, an adequate theory of the process is necessary before any progress can be made in meaningfully interpreting the observed behavior. In short, the "fishing expedition" approach to research is particularly risky when the eye fixation methodology is being used.

Consumer Decision Making

An eye fixation analysis of consumer decision making has recently been completed (Russo, 1977a). Typical shoppers made purchase decisions in a simulated supermarket setting. The subjects' faces were videotaped through a one-way mirror; they were not aware that their eye fixations were being recorded. The results of a typical trial are shown in Figure 2. A schematic diagram of one array of products is shown. The numbers denote eye fixations listed in order of occurrence. The data are hypothetical, but realistic, and are representative of the major characteristics of the 138 actual fixation sequences.

FIGURE 2

REPRESENTATIVE HYPOTHETICAL SEQUENCE OF EYE FIXATIONS FROM A CONSUMER CHOICE TASK

Seneca with Cinnamon 15 oz. 39¢ (29)	Mott's Natural Style 20 oz. 55¢ (28)	S&W Gravenstein 8.5 oz. 27¢	Town House Gravenstein 8.5 oz. 22¢ 2
S&W Gravenstein 15 oz. 43¢ 7, 9, 17	Country Pure unsweetened 14 oz. 43¢ 8, 19, (27)	Town House Gravenstein 15 oz. 39¢ 10, 15	Town House Gravenstein 25 oz. 52¢ 1, 11, 23
S&W Gravenstein 25 oz. 70¢ 6, 18	Mott's 25 oz. 56¢ (26)	Country Pure unsweetened 23 oz. 54¢ 21	Town House Gravenstein 35 oz. 69¢ 12
Seneca 46 oz. 97¢ 5	Apple Time 49 oz. \$1.04 4	Mott's 35 oz. 75¢ ^a 25 3, 14, 16, 20, 22	Town House Gravenstein 50 oz. 99¢ 13, 24

^aThis is the brand/size usually purchased and the one chosen during this decision.

The results of this study will not be summarized here. A few findings, however, serve to illustrate the use of the eye fixation data. The choice process was partitioned into three stages: overview, comparison and checking. The first and last stages were identified primarily by the absence of repeated fixations. In Figure 2, the fixations covered by these stages are: overview, 1-6; comparison, 7-22; and checking, 23-25.

The comparison stage was characterized by pair comparisons. That is, the same alternating pattern of fixations (X-Y-X ...) that was found by Russo and Rosen (1975) occurred during the comparison stage of the consumer decision process. In the preceding example there are three such alternations: 7-8-9, 14-15-16, and 20-21-22. Overall, 50% of the comparison fixations were accounted for by pair comparison alternations. There is an interesting twist connection among these pair comparisons. About two-thirds of them involve the usual brand. Consumers seem to anchor their pair comparisons on their habitual choice. Competing alternatives are directly (pairwise) compared to this standard.

One other finding from this study is relevant to the eye fixation methodology. On over 50% of the trials, the shoppers continued to look at the products on the shelves after they had selected and stated their choice. In Figure 2, the four fixations shown in parentheses, (26) through (29), are typical of this post-decisional pattern of looking. These fixations tend to be located on items that have received little or no attention during the decision process. About 80% of the post-decisional fixations were devoted to brand/sizes never previously looked at or fixated only once before. This is an example of the eye fixation methodology exposing part of the process that would probably not have been revealed by any other method.

Comparison Among Methodologies

In this section, five major methods for identifying cognitive strategies are compared. The methods are evaluated only on their ability to expose cognitive strategies. This evaluation ignores all other experimental goals, such as predicting the output of the process or determining memory structure. First, the five methodologies are described. Then seven performance attributes are presented on which each of the methodologies is evaluated and compared with the others. This entire section is summarized in Table 2, to which readers, if they choose, can turn immediately.

Methodologies

The previous discussion of eye fixations will not be recapitulated here. Instead, we move to the four alternate methodologies that are to be evaluated.

Chronometric Analysis. This method relies upon the total time required to complete a process. This time is typically recorded for many trials within an appropriate experimental design. By comparing the mean response times for different experimental conditions, the nature of a cognitive process can be investigated. This is especially true if the process can be described as a sequence of stages or subprocesses. This technique is most effective for brief tasks, those less than a few seconds in duration. Its use is more problematical for longer tasks like purchase decisions. This technique has seen very little use in consumer research. For an introduction, see Gardner, Mitchell, and Russo (1978); also see Johnson and Russo (1978).

Information Display Boards. In this method, stimulus (product) information is available in an array of pockets on a large board. The subjects reach into a pocket to draw a card which contains the desired information. This method is representative of most techniques that trace a

process through a sequence of overt information acquisition responses. The major exception is eye fixations. The important difference between information display boards and eye fixations is acquisition effort. Whereas an eye fixation typically requires 230 msec. to deliver the requested information (Russo, 1978), a manual reaching response takes about ten times longer.

Input-Output Analysis. This label refers to all methods that rely exclusively on the output of the cognitive process as their data base. By carefully designing the stimulus input, an analysis of the output can reveal processing detail. For example, various heuristics for judgment have been exposed only by analyzing the judgments themselves (Tversky and Kahneman, 1974). See also Lichtenstein and Slovic (1973) and Tversky (1969). Because the method of input-output analysis relies on such a narrow data base, it is not always successful in revealing the underlying process. For example, compare Slovic and MacPhillamy (1974) with Russo and Doshier (1976). Input-output analysis is easily the most common experimental method in behavioral science and should be contrasted with the newer process tracing techniques.

Verbal protocols. A verbal protocol is generated when subjects think aloud while performing an intellectual task. That is, subjects are required to verbalize what they are thinking as they perform the task. This procedure should be differentiated from introspection, in which trained subjects are asked to observe and explain their own internal processes. For a discussion of this distinction, see Payne, Braunstein, and Carroll (1977). The collection of a concurrent verbal protocol should also be distinguished from retrospectively collected protocols. A concurrent protocol avoids many problems associated with protocols that rely on memory (Nisbett and Wilson, 1977). The most important point about verbal protocols is that they are difficult to analyze formally. This is true even for the most structured tasks like logic, chess, and cryptarithmic (Newell and Simon, 1972). They have been successfully analyzed in relatively few consumer experiments (however, see Bettman, 1970; and Payne, 1976).

Performance Attributes

Seven attributes have been selected on which the performance of the five methodologies is to be judged. This list represents the necessary compromise between efficiency and completeness. Others may find some eighth attribute more important than one listed here, especially in a particular application. The seven fall into three broad categories: quality of the data, breadth of applicability of the method, and cost.

Data Quality. The success of any methodology for identifying cognitive strategies depends on its ability to reveal the process in detail. The more detail, the better. A lack of detail must be compensated by repeated trials within complex experimental designs.

If detail roughly corresponds to the quantity of data, then informativeness represents an aspect of data quality. This attribute describes the amount of information conveyed by a single datum, such as the output of a process or one eye fixation. Informativeness also includes some notion of interpretability. The more easily interpretable the data, the more informative.

Just as important as either detail or informativeness is validity. The recorded observations must be a valid indication of the underlying process. They should not be easily censored or distorted as a reaction to the experimental situation. One test for potential invalidity due to reactivity is the question: Can subjects successfully conceal any or all of their strategy from an experimenter who relies on "Method X"?

Breadth of Applicability. To be useful a method should be applicable in a range of experimental settings. These include natural consumer environments, such as the supermarket or at home in front of the TV, as well as laboratory settings.

Breadth of applicability can be limited by the obtrusiveness of the method. Certain situations that are particularly susceptible to experimental reactivity require unobtrusive recording. Would shopping decisions be performed normally if shoppers knew that their eye fixations or total decision time were being recorded?

Costs. The ease of use, or convenience, of a method is important, especially if many trials are to be collected. Note that convenience includes data analysis as well as data collection. Thus, detailed observations like eye fixations are difficult to use, more because of the effort required to analyze them than to collect them.

The last attribute is the cost of equipment, both in dollars and the time necessary to set up an operating system. For example, an eye movement laboratory is costly both because the equipment can be expensive and also because a software system can take considerable time to develop.

Evaluation by Methodology and Attribute

Each of the five methodologies is evaluated on each of the seven performance attributes. A five-category ordinal scale is used: excellent, very good, good, fair, and poor. The values are intended to be consistent within attributes. No comparisons across attributes are implied by the ratings. Relative importance among the attributes depends upon the particular research problem and can be expected to vary greatly. The thirty-five ratings are summarized in Table 2. An explanation of these ratings follows.

Ability to Reveal Process Detail

1. The total duration of the process by itself cannot provide much detail. This is especially true for the relatively long processes associated with typical consumer tasks. In certain circumstances, however, detail can be obtained by performing more sophisticated chronometric analyses on the durations from a complete, prop-

erly designed experiment. For a discussion, see Gardner, Mitchell, and Russo (1978) or reviews in the psychology literature (Chase, 1977; Pachella, 1974).

2. This is the most detailed of the process tracing methods. An interesting comparison with verbal protocols is reported by Newell and Simon (1972, pp. 310ff). They claim that adding eye fixations to their standard analysis of verbal protocols increased the precision of the process trace by about 50%. This comparison is flawed by the reliance of the eye fixation analysis on a prior description of the process derived from the verbal protocols. Further, one of the authors of this study has since claimed that "verbal protocols provide . . . data at densities comparable to the densities of eye fixation records" (Simon, 1976, p. 261). On the other hand, the Newell and Simon data were recorded in a very structured task, cryptarithmic, using highly educated subjects. Both factors enhance the amount of verbal information that can be generated. In contrast, the detail available from eye fixations is present even in less structured tasks, such as viewing advertisements, and even when representative consumers serve as subjects. Overall, the detail provided by verbal protocols is very good, but not as great as that provided by eye fixations.

3. Information display boards reveal less detail than is commonly assumed. Consider a simple measure of detail like the observation rate. This is the total number of observations divided by the total time of the process, expressed in observations/min. The estimate of this rate for the information display board method is derived from data reported by Jacoby and Chestnut (1977). For comparison purposes, the rate for eye fixations is obtained from Russo (1977a). This latter study is similar to Jacoby and Chestnut's. They both used representative shoppers making realistic purchasing decisions from an array of real brands. However, there were differences in the display presented to the subjects. Whereas the stimulus for Jacoby and Chestnut was an information display board (brand and attribute array), Russo used actual product packages on shelving that simulated a supermarket setting. The reported mean observation rates are 4-5 acquisitions/min with the information display boards¹ versus 49 eye fixations/min. Thus,

¹The lower bound of four acquisitions/min may be high. This value was based only on the 61% of the subjects who

TABLE 2

A COMPARISON OF FIVE METHODOLOGIES FOR CONSUMER INFORMATION PROCESSING RESEARCH

Method	Performance Attributes						
	Data Quality			Breadth of Applicability		Cost	
	Detail Revealed	Informativeness	Validity	Range of Settings	Unobtrusiveness	Ease of Use	Equipment Price
Chronometric Analysis	1. poor	6. good	11. v. g.	16. good	21. fair	26. v. g.	31. good
Eye Fixations	2. exc.	7. fair	12. exc.	17. fair	22. good	27. fair	32. poor
Information Boards	3. fair	8. poor	13. good	18. v. g.	23. fair	28. exc.	33. exc.
Input-Output Analysis	4. poor	9. v. g.	14. good	19. exc.	24. v. g.	29. exc.	34. exc.
Verbal Protocols	5. v. g.	10. exc.	15. fair	20. exc.	25. poor	30. good	35. v. g.

the density of observational detail is almost ten times less for information boards than for eye fixations.

Van Raaij (1976b) used an information acquisition response that is intermediate in effort between a reaching response and an eye fixation. His subjects reached for an actual product package which displayed four information attributes on the four sides. Compared with an information display board, Van Raaij's reaching response is no more effortful and possibly less so. And without any further reaching, the subjects could acquire three additional attributes by turning the package in their hands. Given this level of acquisition effort, one would expect an observation rate between those of information display boards (4-5 observations/min) and eye fixations (49 observations/min). The value calculated from Van Raaij's data is 14 observations/min.

Because a reaching response is relatively effortful, subjects can be expected to adopt strategies that minimize the use of this type of information acquisition. For example, rather than reacquire the same value a second or third time, they tend to memorize these values. When needed again they are recalled, which is much quicker than a second reaching response. The problem is that, since recall is internal, it is unobservable. This defeats the purpose of the method, to provide a detailed trace of the choice process. In the information display board studies by Jacoby and his associates, reacquisition rate never exceeds 7% of all acquisitions (Jacoby and Chestnut, 1977; Jacoby, Chestnut, Fisher, and Weigl, 1976). When Van Raaij used the information board method he found a zero reacquisition rate (Van Raaij, 1977). In contrast to these low values, the reacquisition rate for eye fixations is much higher. Russo (1977a) reports a rate of 56%; i.e., 56% of all eye fixations were re-fixations. Van Raaij (1976a) also recorded eye fixations but required a simple, manual manipulation of packages. For this information acquisition response, the reacquisition response was 35%. As expected, this value is intermediate between the information display board rate (< 7%) and the pure eye fixation rate (56%). Caution must be used in interpreting these data because differences in the stimulus display make reacquisition rates not precisely comparable. However, such large differences indicate that at least some difference would remain if identical stimulus display could be used across all methods.

The danger in all this is that because information display boards yield a relatively sparse trace of the process, they may lead to different views of the subjects' strategies. Using a detailed eye movement trace, both Russo (1977a) and Van Raaij (1976a) report that subjects use pair comparisons to directly compare among alternatives. In contrast, Jacoby and Chestnut (1977, p. 94) conclude that their subjects' information "search seems to be of very shallow depth." Arch, Bettman, and Kakkar (1978), use verbal protocols to provide insight into information display board behavior. They report that "there seems to be a great deal of exploratory or curiosity-based information acquisition in the display board task." These contrasting inferences from the two methodologies suggest that the danger of sparse data yielding different conclusions is real.

In summary, information display boards require an effortful acquisition response, more effortful than the eye fixations on which shoppers mainly rely in natural settings. This increased effort encourages the use of strategies that conceal much of the detail of information acquisition.

searched the information board at all (when brand name was given). The data reported by Jacoby and Chestnut (1977, Tables 6-10) do not permit a revised calculation of observation rate based on all subjects.

4. Input-output analysis was not intended to provide detailed observation of the process. This method relies entirely on one observation, the output of the process. Only with repeated observations in an appropriate experimental design can a cognitive strategy be inferred. For a notably successful instance, see Tversky and Kahneman (1974). However, important parts of the strategy, if not its entire detail, can sometimes be exposed; see, e.g., Slovic, Fleissner, and Bauman (1972) or Slovic (1969).

5. Verbal protocols reveal more detail than any method except eye fixations. Although quantitative comparisons are problematical, the observation rates of the two methods can be compared, just as they were for eye fixations and information display boards. Newell and Simon (1972, pp. 310ff) formally analyze a verbal protocol, transforming it into a series of "productions." These are the smallest meaningful cognitive units within their theory. They report a rate of 19 productions/min. When the associated eye fixations are analyzed, in conjunction with the verbal protocol, the rate increases to 28 productions/min. This comparison is biased in favor of eye fixations because they have the benefit of a prior verbal analysis. However, it has the rare advantage of comparing two different types of data on the same criterion.

Verbal protocols are difficult to compare to eye fixations because these two types of data are so different. The strength of verbal protocols is their ability to reveal larger strategic concepts. These larger concepts are easier to verbalize than details like the purpose of an individual eye fixation. Subjects do not verbalize many details either because they are not aware of them or reporting them would slow down performance intolerably. An indication of how much more detail is available than is spoken can be obtained from results reported in Russo (1977b). A standard concurrent verbal protocol was compared to a new technique called prompted protocols. A retrospective protocol is collected while subjects are prompted by a replay of their sequence of eye fixations. Based on three different tasks, the prompted protocols averaged over twice as many words as the concurrent protocols. When coded for specific informative statements, the prompted method yielded about 75% more statements than a concurrent protocol. This additional information in a prompted protocol illustrates the details that are otherwise not reported, but would seem to be captured by an eye fixation sequence.

In summary, in spite of the difficulty of directly comparing verbal protocols with nonverbal data, the evidence consistently suggests that verbal protocols do not provide as detailed a process trace as eye fixations. Nonetheless, verbal protocols still provide very good detail and are clearly superior to the other three methods on this attribute.

Informativeness

6. Response times are directly interpretable and are, in this sense, informative. It may take many of them to provide a complete picture of the underlying process, but each response time is still clearly informative.

7. Eye fixations are inherently informative only in so much as characteristic patterns of fixations can be isolated. The individual fixations are usually not interpretable. Of course, with an adequate theory of the process, an eye fixation sequence can become very informative, but the existence of such a theory is not presumed for the typical research problem.

8. The sequence of acquisition responses provided by the information display board method suffers from the same difficulties as eye fixations. In addition, because the response rate is lower, there is less oppor-

20. A major advantage of verbal protocols is that they can be collected in almost any setting. It is almost as easy to record shoppers talking in a supermarket (e.g., Bettman, 1970) as in a controlled laboratory situation. Because a relaxed, comfortable subject generates better protocols, it can be argued that the method works better in natural field settings than in the laboratory.

Unobtrusiveness

21. Because of the need for speed-accuracy instructions and appropriate measuring apparatus, chronometric methods tend to be very obtrusive. It is difficult to insure uniform task performance without making subjects aware that their response times are being recorded. However, unobtrusive recording is possible, usually by hidden observers using stopwatches.

22. Eye fixations can be recorded unobtrusively (e.g., Russo, 1977a; Van Raaij, 1976a; Monty, 1975). However, the experimenter must pay a price in either more expensive equipment or more costly manual analysis.

23. Information boards are necessarily obtrusive. The special stimulus display means that subjects know their behavior is being observed. On the other hand, it ought to be possible to construct a display very similar to an information board that would simulate the information array in a natural setting. Unobtrusive observation of information acquisition should then be possible. This author is not aware of any such studies.

24. Unobtrusive recording is easiest for the input-output method. The only response that has to be recorded is the output. However, when used for process analysis, a highly structured experimental design is usually needed. This design can be difficult to conceal from subjects.

25. There is no way to ask subjects to think aloud without obtruding into their natural behavior. There are just not enough subjects who normally go around talking to themselves.

A major distinction between verbal protocols and the other four methodologies is that only the former requires an extra response. Subjects must simultaneously perform a second task, to generate a coherent verbal description of the thoughts that occur to them while completing the primary task. The existence of this second task is not only obtrusive, but raises the larger question: does generating a verbal protocol alter the primary process in important ways? Most experimenters believe that the answer is negative; however, there is little, if any, direct empirical support for this conclusion.

Ease of Use

26. The recording of response times is relatively easy, although standardized procedures should be followed (see Sidowsky, 1966). The data analysis is similarly uncomplicated.

27. Some eye fixation methods are easy to use, but most require either complicated recording procedures or effortful data analysis. Unlike the other methodologies, there is such a wide range of techniques for recording eye fixations that one can almost always trade off such desiderata as accuracy, ease of use, and unobtrusiveness.

28. A major advantage of information boards is that they are very easy to use. Large numbers of subjects can be conveniently run without special selection or training (see Jacoby and Chestnut, 1977).

29. In special situations, recording the output of a process can become complicated. In general, however, the

input-output method is very easy to use.

30. Verbal protocols are considered relatively easy to use, because a verbal report is such an easily accessible response. This impression is too optimistic for two reasons. As many a well-intentioned experimenter has found out, analyzing the protocols can prove problematical. A formal analysis is effortful and probably restricted to highly structured tasks (see Newell and Simon, 1972). And even "formal" analyses require enough subjective interpretations, that marked disagreement among data analysts is possible (Haines, 1974). A second problem is that many subjects, especially those with less education, do not generate an acceptable protocol when requested. The classic protocol studies of Newell and Simon use graduate students as subjects. This writer's experience is that typical shoppers provide sparse, self-conscious protocols. A training procedure is required before representative consumers can generate verbal protocols that are informative.

Equipment Price

31. For the accurate control of stimulus exposure and recording of response duration, relatively expensive tachistoscopes are normally used. Alternatively, a laboratory computer can control shutters or a CRT display. Less expensive recording devices can trade off cost for accuracy. Even a simple stopwatch can be very effective.

32. Basically, eye fixation equipment is expensive. Furthermore, it often requires other equipment, like a laboratory computer, or special facilities, like light-sealed rooms with electrical shielding. On the other hand, the eager but impoverished researcher can use simple video recording or can construct his own device for only a few hundred dollars (e.g., Russo and Mathews, 1975). It is hoped that the example discussed earlier (Russo, 1977a) will expose how low the financial barrier to recording eye fixations can be.

33. Information boards are inexpensive, durable, and easy to maintain. This can be especially advantageous to the researcher who wants to trace information acquisition behavior without making the commitment required by eye fixation technology.

34. The equipment needed to display stimuli and to record output varies with the nature of the task. Usually, however, the cost is minimal.

35. To collect verbal protocols, a good microphone and an adequate audio recorder are needed, with a transcriber for typing. Although this equipment can cost several hundred dollars, it may already be available to most consumer researchers or can be rented for a small fee.

Joint Methodologies

A catalogue of the strengths and weaknesses of individual methodologies leads to the consideration of the concurrent use of two or more methodologies. There are two modes of joint use: parallel and interactive. Methods are used in parallel when two sets of observations are recorded and neither set of data is changed from the situations where it alone is observed. In the interactive situation, the observations themselves are determined by a joint methodology. They differ from the data generated by either single method. Both cases are considered below.

Parallel Joint Methodologies

Parallel joint methodologies are a natural way to minimize the shortcomings of one method. For instance, where eye fixations have only a fair rating on informativeness, verbal protocols are excellent. In the con-

tunity for patterns of responses to occur and to repeat themselves. Because of a preference for recall over re-acquisition, stimulus components are seldom viewed more than once. This further inhibits the development of characteristic patterns of responding.

9. The informativeness of the output of a process is usually very good. A shopper's chosen brand, the rated importance of an information source, or the amount of product knowledge that can be recalled are all informative and easily interpretable responses. Of course, a single output response cannot inform the entire process; but a suitable pattern can.

10. Verbal protocols provide more information per datum than any of the other methods. Indeed, their preeminent advantage is that they usually provide their own interpretation. That is, at the same time that verbal protocols trace the process, they also explain it. This self-interpretability is unique. Self-interpretability is particularly advantageous when the research is exploratory, i.e., when there is no well-founded theory to guide the investigation of the process of interest.

Validity of the Recorded Observations

11. Chronometric analysis is susceptible to two threats to validity, a trading-off of accuracy for speed and a heightened level of motivation. In order to assure uniformity across all trials, subjects are typically given explicit instructions about speed and accuracy. This means that they are made aware that their performance is being timed. Such awareness may cause faster or more consistent performance than might be the case if subjects could be observed unobtrusively.

12. The validity of eye fixations is excellent. Since people are typically unaware of these movements, they are difficult to alter or censor. Because of the apparatus usually employed to measure eye fixations, reactivity is a potential problem. However, the evidence suggests that subjects soon forget that their eye fixations are being recorded, i.e., that eye fixation recording does not alter the ongoing cognitive process (e.g., Gilbert and Gilbert, 1942). If reactivity is still worrisome, there are methods for unobtrusively recording eye fixations (e.g., Van Raaij, 1976a; Russo, 1977a).

13. The validity of the information board methodology is problematical and unknown. On the face of it, one would expect it to be very good. However, total reliance on reaching response is unrealistic and this may alter at least some aspects of the cognitive process (e.g., the rate of repeated acquisition). It is also worrisome that the stimulus environment is so different from natural shopping settings. Clearly, what is needed is a validity check on the method. Ideally, the information board results should be compared with information acquisition in natural shopping settings.

Van Raaij (1977) reports results that approximate this ideal comparison. The same shoppers chose in the same product category using both naturalistic packages and an information display board for the identical package information. Van Raaij reports marked differences in the observed behavior, including the total amount of information acquired and the reacquisition rate. The number of brands considered averaged 10.4 of 13 when packages could be picked up and examined, but only 6.6 of 13 when subjects had to use an information display board. These differences suggest that subjects' behavior in the information board situation may differ in important ways from that in a natural shopping setting.

14. The validity of the process output is good. This method does not receive a higher rating because it is usually possible for the subjects to alter their output

(and the strategies that generate them). Social desirability or just reactivity to the experimental situation can change the process. Of course, if the output is recorded unobtrusively, the validity of the method is excellent. The problem is that a complex experimental design is almost essential if input-output analysis is to expose a cognitive strategy. Such structured designs are difficult to conceal from subjects.

15. Recall the test: "Can subjects successfully conceal any or all of their strategy from the experimenter who relies on Method X?" Applying this test to verbal protocols results in no better than a "good" rating for validity. Of course, the problem is worse if the protocols are retrospective (Nisbett and Wilson, 1977). In this case, a "fair" rating is in order; and experimenters should try to construct a validity check. Even with concurrent protocols, however, there is an irresistible tendency for subjects to "clean up their act," i.e., to describe a more coherent or thorough strategy than normally occurs. In addition, subjects are not aware of many details, such as quick checks, sidetracks, brief interruptions, etc. These behaviors are not included in the verbal protocol, further creating an overly positive picture of the cognitive process.

Range of Experimental Settings

16. Because of the need for specific speed-accuracy instructions and for accurate measuring equipment (reaction time clocks, tachistoscopes, etc.), chronometric methods tend to be restricted to laboratory situations. Response times can be recorded in the field, although simpler, less accurate apparatus must be used.

17. Eye fixations are the least transportable of all the methodologies. Only if video recording is used (which requires a lengthy hand analysis) can they be used in the field. Eye fixations have been recorded from shoppers in a simulated supermarket setting (Russo, 1977a). However, this writer knows of no recordings of actual supermarket behavior. (This could be accomplished with one or more small concealed video cameras with wide angle lenses, but considerable effort would be required to analyze the data.)

A portable eye fixation recording apparatus has been developed and used extensively by Mackworth (see Thomas, 1968). Subjects wear a helmet that carries both the eye position sensor and a video camera. The camera records the scene in front of the subject (i.e., the camera "sees" what the subject sees). The apparatus then superimposes on the video recording a bright spot that indicates the point of regard within the visual scene. This apparatus is not only transportable, but permits the use of dynamic stimulus displays, such as moving scenes, TV advertisements, etc.

Overall, the range of settings where eye fixations can be recorded is almost unlimited. The relatively low rating for this methodology reflects the difficulty of attaining this flexibility. This difficulty translates into a greater cost of equipment or the increased difficulty of data analysis.

18. Information display boards are easily transportable. Jacoby has used them in a specially designed trailer parked in a shopping center for easy recruitment of consumer subjects (Jacoby and Chestnut, 1977). This method will always be somewhat laboratory-bound, however, because it requires a special stimulus display. This is the only method that does not permit naturalistic product displays.

19. From the viewpoint of its technology (apparatus, etc.), this is the simplest method. It is very easy to transport to field situations.

sumer choice study reported earlier (Russo, 1977a) it was not clear that the eye fixations would be easily interpretable, so a concurrent verbal protocol was collected. These data were used to partition the process into stages (overview, comparison, and checking). In addition, they provided insight into the specific product attributes that subjects considered (e.g., price, size, ingredients, packaging, etc.). It was not possible to identify the relevant attribute(s) from the eye fixations on a package.

More complex use of several methods is entirely feasible. Russo and Doshier (1976) explore the binary choice process by using eye fixations, verbal protocols, input-output analysis, and response times. Payne and Braunstein (1977) use verbal protocols, response times, and a computer-controlled information acquisition procedure that resembles information display boards. For a further discussion of parallel joint methodologies, see Payne, Braunstein, and Carroll (1977).

Interactive Joint Methodologies

An interactive joint methodology is as much the creation of a new method as it is the joint use of two standard ones. For example, Reder (1973), Rayner (1975), and Just and Carpenter (1976b) have all developed systems for changing a visual display contingent on where a subject is looking. This is an interactive joint use of input-output analysis and eye fixations. Another example is prompted protocols (Russo and Rosen, 1975; Russo, 1977b). A retrospective verbal protocol is prompted by a replay of the sequence of eye fixations. The verbal protocol can be prompted by other behavior, including the sequence of acquisitions from an information display board (Arch, Bettman, and Kakkar, 1978). Prompted protocols are an example of the interactive, joint use of eye fixations (or information display boards) and verbal protocols.

The possibilities for interactive joint methodologies are almost limitless. This is an exciting research area, both for solving current data acquisition problems and also for generating new forms of data that will yield insight never before possible.

Summary and Conclusions

The eye fixation methodology has important advantages not found in other methods. These process tracing data offer excellent detail and validity. The drawbacks to this method are primarily practical considerations like cost, transportability, and ease of use. Although these drawbacks cannot be ignored, there are trade-offs that can permit the recording of eye fixations at low cost or in natural settings or unobtrusively. The main advantage of this method, however, is the high quality of the data, which is, after all, the best reason for using any research method.

Verbal protocols are remarkably complementary with eye fixations. An examination of Table 2 reveals that no method is more different from eye fixations than verbal protocols. This suggests that the disadvantages of one can be compensated by the strengths of the other. Prompted protocols are one illustration of this. They preserve the informativeness of verbal protocols while adding the detail of eye fixations. In general, it is suggested that any researcher using eye fixations would benefit from the simultaneous use of verbal protocols. These methods naturally complement each other.

This review has been less sanguine about the use of information display boards. In several respects they generate data of low quality that seem to differ from that found in natural situations. The problem is the high effort associated with a reaching response. If this effort could be reduced, methods very similar to infor-

mation display boards could provide adequate data quality at low cost. A computer-controlled information display is one way in which this might be accomplished (Payne and Braunstein, 1977).

The main conclusion of this comparison of methods is that no one method dominates. Each one has its own set of advantages. The choice of a method depends on the specific experimental situation. Generally, the use of joint methodologies can make up for the deficiencies in one method. If this paper yields one firm recommendation, it is that researchers make greater use of the opportunities and advantages of joint methodologies.

REFERENCES

- Mathew Alpern, "Movements of the Eyes," in H. Davson (ed.), The Eye, Vol. 3 (New York: Academic Press, 1969).
- David C. Arch, James R. Bettman and Pradeep Kakkar, "Subjects' Information Processing in Information Processing in Information Display Board Studies," in H. Keith Hunt (ed.), Advances in Consumer Research, Vol. 5.
- James R. Bettman, "Information Processing Models of Consumer Behavior," Journal of Marketing Research, 7 (1970), 370-376.
- William G. Chase, "Elementary Information Processes," in W. K. Estes (ed.), Handbook of Learning and Cognitive Processes, Vol. 5 (Hillsdale, N. J.: Lawrence Erlbaum Associates, in press).
- John S. Conery, R. T. Smith and J. Edward Russo, "Interactive Computer Techniques," Proceedings of the Digital Equipment Users Society, 1 (1975), 389-391.
- Hugh Davson (Ed.), The Eye, Vol. 1-6 (New York: Academic Press, 1962-1975).
- Meryl P. Gardner, Andrew A. Mitchell and J. Edward Russo, "Chronometric Analysis: An Introduction and an Application to Low Involvement Perception of Advertisements," in H. Keith Hunt (ed.) Advances in Consumer Research, Vol. 5 (Association for Consumer Research, 1978, forthcoming).
- A. J. Gilbert and D. W. Gilbert, "Reading Before the Eye-Movement Camera Versus Reading Away from It," Elementary School Journal, 42 (1942), 443-447.
- Frank M. Goode and J. Edward Russo, "Computerized System for Collecting Eye Movement Data," Proceedings of the 78th Convention of the American Psychological Association (1970), 45-46.
- George H. Haines, "Process Models of Consumer Decision Making," in G. D. Hughes and M. L. Ray (eds.), Buyer/Consumer Information Processing (Chapel Hill, N. C.: University of North Carolina Press, 1974).
- Jacob Jacoby and Robert W. Chestnut, "Amount, Type, and Order of Package Information Acquisition in Purchasing Decisions," Final Report to National Science Foundation (GI-43687), 1977.
- Jacob Jacoby, Robert W. Chestnut, William A. Fisher and Karl C. Weigl, "Simulating Nondurable Purchase: Individual Differences and Information Acquisition Behavior," Purdue Papers in Consumer Psychology, No. 153, 1976.
- Eric J. Johnson and J. Edward Russo, "The Organization of Product Information in Memory Identified by Recall Time," in H. Keith Hunt (ed.), Advances in Consumer Research, Vol. 5.

- Marcel A. Just and Patricia A. Carpenter, "Eye Fixations and Cognitive Psychology," Cognitive Psychology, 8 (October 1976), 441-480, a.
- Marcel A. Just and Patricia A. Carpenter, "The Role of Eye Fixation Research in Cognitive Psychology," Behavior Research Methods and Instrumentation, 8 (1976), 139-143, b.
- Paul A. Kollers, "Buswell's Discoveries," in Eye Movements and the Higher Psychological Functions.
- Sarah Lichtenstein and Paul Slovic, "Response-Induced Reversals of Preference in Gambling: An Extended Replication in Las Vegas," Journal of Experimental Psychology, 101 (1973), 16-20.
- Geoffrey Loftus, Paul Mathews, Susan Bell and Stephen Poltrack "General Software for an On-Line Eye Movement Recording System," Behavior Research Methods and Instrumentation, 6 (1974), 525-530.
- Norman H. Mackworth and E. Llewellyn Thomas, "Head-Mounted Eye-Marker Camera," Journal of the Optical Society of America, 52 (1962), 713-716.
- Richard A. Monty, "An Advanced Eye-Movement and Recording System Featuring Unobtrusive Monitoring and Automatic Data Processing," American Psychologist, 30 (1975), 331-335.
- Richard A. Monty and John W. Senders (eds.), Eye Movements and Psychological Processes (Hillsdale, N. J.: Lawrence Erlbaum Associates, 1976).
- Alan Newell and Herbert A. Simon, Human Problem Solving (Englewood Cliffs, N. J.: Prentice-Hall, 1972).
- Robert G. Pachella, "The Interpretation of Reaction Time in Information-Processing Research," in B. H. Kantowitz (ed.), Human Information Processing: Tutorials in Performance and Cognition (Hillsdale, N.J.: Lawrence Erlbaum Associates, 1974).
- John W. Payne, "Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis," Organizational Behavior and Human Performance, 16 (1976), 366-387.
- John W. Payne and Myron L. Braunstein, "Contingent Processing in Risky Choice: A Process Tracing Investigation," unpublished manuscript, University of Chicago, 1977.
- John W. Payne, Myron L. Braunstein and John S. Carroll, "Exploring Pre-Decisional Behavior: An Alternative Approach to Decision Research," unpublished manuscript, University of Chicago, 1977.
- Steven M. Reder, "On-Line Monitoring of Eye Position Signals in Contingent and Non-Contingent Paradigms," Behavior Research Methods and Instrumentation, 5 (1973) 218-228.
- J. Edward Russo, "An Eye Fixation Analysis of Consumer Shopping Decisions" (tentative title), in preparation, 1977, a.
- J. Edward Russo, "Verbal Protocols Prompted by Eye Fixations" (tentative title), in preparation, 1977, b.
- J. Edward Russo, "Adaptation of Cognitive Processes to the Eye Movement System," in Eye Movements and the Higher Psychological Functions.
- J. Edward Russo and Barbara Anne Doshier, "An Information Processing Analysis of Binary Choice," unpublished manuscript, Carnegie-Mellon University, 1976.
- J. Edward Russo and Paul Mathews, "Tracking Eye Fixations in Cognitive Tasks," unpublished manuscript, University of California, San Diego, 1975.
- J. Edward Russo and Larry D. Rosen, "An Eye Fixation Analysis of Multialternative Choice," Memory and Cognition, 3 (May 1975), 267-276.
- John W. Senders, Dennis F. Fisher and Richard A. Monty (eds.), Eye Movements and the Higher Psychological Functions (Hillsdale, N. J.: Lawrence Erlbaum Associates, 1978).
- Joseph B. Sidowski (ed.), Experimental Methods and Instrumentation in Psychology (New York: McGraw-Hill, (1966).
- Herbert A. Simon, "Discussion: Cognition and Social Behavior," in J. S. Carroll and J. W. Payne (eds.), Cognition and Social Behavior (Hillsdale, N. J.: Lawrence Erlbaum Associates, 1976).
- Paul Slovic, "Analyzing the Expert Judge: A Descriptive Study of a Stockbroker's Decision Process," Journal of Applied Psychology, 53 (August 1969), 255-263.
- Paul Slovic and Douglas MacPhillamy, "Dimensional Commensurability and Cue Utilization in Comparative Judgment," Organizational Behavior and Human Performance, 11 *k864(m k62-194).
- Paul Slovic, Dan Fleissner and W. Scott Bauman, "Analyzing the Use of Information in Investment Decision Making: A Methodological Proposal," The Journal of Business of the University of Chicago, 45 (April 1972), 283-301.
- E. Llewellyn Thomas, "Movements of the Eye," Scientific American, 219 (1968), 88-95.
- Amos Tversky, "Intransitivity of Preferences," Psychological Review, 76 (1969), 31-48.
- Amos Tversky and Daniel Kahneman, "Judgment under Uncertainty: Heuristics and Biases," Science, 185 (1964), 1124-1131.
- W. Fred Van Raaij, "Direct Monitoring of Consumer Information Processing by Eye Movement Recorder," Tilburg Papers on Consumer Evaluation Process, No. 12, 1976a.
- W. Fred Van Raaij, "Direct Observation of Consumer Information Processing," Tilburg Papers on Consumer Evaluation Processes, No. 11, 1976b.
- W. Fred Van Raiij, "Consumer Information Processing for Different Information Structures and Formats," in W. D. Perreault, Jr. (ed.), Advances in Consumer Research, Vol. 4 (Association for Consumer Research, 1977).
- Lawrence R. Young and David Sheena, "Survey of Eye Movement Recording Methods," Behavior Research Methods and Instrumentation, 7 (1975), 397-429, a.
- Lawrence R. Young and David Sheena, "Eye Movement Measurement Techniques," American Psychologist, 30 (1975), 315-330, b.

VERBAL PROTOCOLS AND DIRECT OBSERVATION
OF SUPERMARKET SHOPPING BEHAVIOR: SOME
FINDINGS AND A DISCUSSION OF METHODS

John W. Payne, Duke University
E.K. Easton Ragsdale, University of Chicago

Abstract

A prime characteristic of the information processing approach to the study of behavior is the use of methods that are capable of providing detailed observations of the course of behavior over time (Simon, 1976). Examples of such process tracing techniques are: the collection of verbal protocols, monitoring information acquisition, and response time recording. The purpose of this paper is to discuss some of the benefits and some of the problems associated with the use of one process tracing technique--verbal protocols, to investigate consumer decision behavior. In particular, the use of verbal protocols to study consumer decision making in supermarkets will be illustrated.

Verbal Protocols

The collection of verbal protocols is conceptually a straightforward method of obtaining process data. The subject is simply asked to give continuous verbal reports, "to think aloud," while performing the task of interest. The verbal protocol is then treated as a "record of the subject's ongoing behavior, and an utterance at time t is taken to indicate knowledge or operation at time t " (Newell & Simon, 1972).

Distinguishing Protocols from Other Verbal Data

The collection of verbal reports is an old idea in psychology. It is also an idea which has generated considerable controversy. Because there is an apparent parallel between verbal protocols and other types of verbal data, I want to stress a few of the distinctive features of the process tracing technique of verbal protocols. First, unlike the earliest introspective techniques, subjects under the process tracing strategy are naive about the theoretical constructs of interest to the researcher. Second, in collecting verbal protocol data the subject is not asked to theorize about his behavior. Instead the subject is asked "only to report the information and intentions that are within his current sphere of conscious awareness" (Newell & Simon, 1972). The researcher, not the subject, does all the theorizing about the cause and consequences of the subject's knowledge state. Third, unlike the traditional way in which verbal reports have been collected--after the response, verbal protocols are collected during the actual performance of the task. While answers to the questions, "What are you doing right now?" may be misleading, they are likely to be considerably less misleading than later descriptions of how the task was performed. Later descriptions allow much more opportunity for the subject to mix current knowledge with past knowledge, making reliable references from verbal reports difficult.

This last distinction is crucial, for the controversy of what subjects can verbally report has reemerged many times. Nisbett and Wilson (1977) recently argued that people have little or no ability to directly observe and verbally report upon higher order mental operations that result in a response of some kind. While the research they review is impressive in supporting their argument, the fact that the evidence is derived entirely

from verbal reports collected after the response limits its relevance for evaluating the process tracing technique of verbal protocols. On the other hand, the evidence from years of research on human problem solving behavior does indicate that verbal protocol data can serve as a valuable basis for model building. Obviously, more research is needed to determine the levels and amounts of information about cognitive processes that can be provided by verbal protocols collected in various task environments.

Using Protocol Data

Once the verbal protocols have been collected, they can be used in a variety of methods. For example, verbal protocols may be a particularly valuable source of data during the early phases of an investigation of consumer behavior. The data can be used to search for regularities in behavior. In addition, a preliminary coding of the protocols may serve as the basis for the calculation of descriptive statistics relevant to theoretical issues. Svenson (Note 1), for example, had six subjects make hypothetical decisions to buy one of seven houses presented in great detail in booklets, one for each house. This study took place in a controlled laboratory environment. Information was obtained on the attributes being used to judge the houses. Information was also obtained on the evaluation of a particular house on a particular attribute in terms of a comparison with a criterion or in terms of a comparison with another alternative. Finally, Svenson was able to track possible changes in decision rules over time. The study of supermarket shopping behavior to be presented later represents an example of the use of verbal protocols to do exploratory consumer research in field settings.

Another very valuable use of verbal protocols is to confirm and extend the interpretations of data collected by other methods. In particular, verbal protocols have already been used successfully to support the interpretation of process data collected by the monitoring of information acquisition behavior (Payne, 1976). In addition, Russo and Doshier (Note 2) have used a "prompted protocols" for the same purpose. Prompted protocols are verbal reports given by the subject after the decision has been made but prompted by presenting the subject with a record of his behavior during the task.

Verbal protocols can also be used to test specific hypotheses about behavior in experimental settings. An example of this use of verbal protocols is provided by Montgomery (1976). He was interested in replicating Tversky's (1969) study on intransitive preference patterns. However, unlike Tversky's procedure, the subjects in Montgomery's study were asked to "think aloud" while making their choices. Statements were coded with the focus on explicit dimensional comparisons such as "fairly small differences" or "they do not differ so much." The coded protocols were used to test theoretical predictions concerning whether a lexicographic choice rule or an addition of utility differences procedure was employed by the subjects. Finally,

verbal protocols have been used by a number of researchers as a basis for computer models of consumer behavior (e.g. Bettman, 1970; Haines, 1974; Payne, 1976). A good discussion of the prospects and problems of such models may be found in Bettman (1974). A more extensive discussion of verbal protocols and other process tracing methods is provided in Payne, Braunstein, and Carroll (Note 3).

The next section of the paper will illustrate the use of verbal protocols, and other more traditional marketing tools (questionnaires, consumer purchase diaries), to study the behavior of consumers shopping in supermarkets.

Supermarket Shopping Behavior¹

Method

The subjects were a convenience sample of 19 housewives with children in the five to twelve year old age bracket. The subjects all lived in Chicago area suburbs. The firm sponsoring the study, a candy manufacturer, set the condition that the housewives have school-age children; previous marketing research indicated that consumers meeting this requirement were most likely to be heavy purchasers of candy. Shoppers were recruited by offering a \$30 payment to each individual who completed all the study requirements. Each shopper also received a small gift part way through the study.

Each shopper supplied verbal protocols during six "major" weekly shopping trips to the supermarket. The shoppers were accompanied by female graduate students acting as observers. At the beginning of each trip, the observers read a standard set of instructions to the subjects:

"We want you to think aloud while shopping. Say what you are doing and where you are going in the store. Whenever you see something as you are walking through the store, mention it. Talk about the products and brands that you are buying as well as those you notice but do not buy. For instance, say what observations you are making about brands, their location in the store, package size, price, ingredients, flavors, and so on. Remember, say anything that comes to your mind no matter how unimportant it seems to you."

When necessary, the observers also made use of a set of standardized prompts to remind the subjects to continuously think aloud:

1. "Remember, say anything that comes to your mind, no matter how unimportant it seems to you."
2. "Say what you are doing now."
3. "Remember talk about whatever you see."
4. "Talk about what observations you are making."
5. "Mention anything that you are thinking about."
6. "Be sure to say what you are doing now."

Great care was taken to minimize the amount of prompting and the amount of interaction between observer and subject.

Finally, in order to familiarize the subjects with the procedure of thinking aloud, they were asked to give verbal protocols while examining a small booklet containing full page magazine advertisements. This task

was completed at the subject's home prior to the first shopping trip.

A great deal of information was acquired in addition to the verbal protocol data. Demographic and shopping behavior data were obtained from a small questionnaire completed by the subjects when they were recruited. Consumer purchase diaries were maintained during the six-week period of in-store observation and during a two-week period either immediately preceding or following the observation period. While in the store, the observers recorded the following information about the shopper's behavior: (1) the path taken through the store; (2) any items taken from major point-of-purchase displays (free standing or end-of aisle displays); (3) the use of a shopping list (all shopping lists were collected by the observers); (4) the use of coupons; and (5) the shopper's nonverbal behavior when passing the candy department.

The methods used in our study extend the previous efforts to study shopping behavior using protocols along several dimensions (e.g., Bettman, 1970; King, 1969; Haines, 1974). For example, Bettman (1970), in perhaps the best known protocol study in marketing, used protocols obtained from only two shoppers. The protocols were used to develop a model of each subject in the form of a decision net or tree. Our study utilized a much larger sample of shoppers. In addition, protocols were obtained over a large number of shopping trips. Also, unlike the Bettman study, the subjects in the present study were not known personally by the researchers, and interactions were kept to a minimum during the shopping trip, e.g., use of only a structured set of prompts. We also did not attempt to build an idiosyncratic model of each subject. Finally, our use of a variety of data collection methods enabled us to cross validate our results. We strongly believe that a multi-method approach is essential to the study of consumer behavior processes.

Results

The final four verbal protocols for each shopper were treated as the basic data base. The first two shopping trips were considered to be practice trials.

The first step in the protocol analysis was to make a complete transcript of the verbal reports given by each shopper. The next step was to break the protocols up into single statements or phrases. According to Newell and Simon (1972), each phrase should correspond to a naive assessment of what constitutes a single task assertion or reference by the subject. Breaking up protocols in this way serves to clarify the sequential nature of the observations and to create units of analysis. As an example of this procedure, excerpts from one protocol are presented below. Complete transcripts can be obtained from the second author.

- D1: OK, alongside the window, we have Heritage House mayonnaise.
D2: and it is on sale this week.
D3: and it is only 79 cents;
D4: and I'm going to take one quart.
D5: Now I am going to buy tuna fish,
D6: that is Starkist tuna fish;
D7: and it is 49 cents a can;
D8: I am going to buy six cans;
D9: OK, 2, 4, 6 Starkist tuna fish, OK.
D10: Along the window they have the specials
D11: Welch grape juice
D12: Oh, they have the great big bottles.
D13: \$1.09.
D14: I did not know that they had the big ones.
. . .

¹The research reported in this section of the paper was conducted by Robert Blattberg, Harry Davis, and the two authors, at the University of Chicago.

D152: Now, there is the apple juice
 D153: They like that
 D154: Hawaiian Punch
 D155: it is \$2.47
 D156: usually they have the orange juice over here
 D157: tea bags
 D158: They have their own brand
 D159: \$1.39
 D160: I bought National tea bags
 D161: and they were pretty good
 D162: I kind of like them
 D163: so I am not about to switch
 .
 .
 D429: Let us look at some of this zucchini squash,
 D430: and then that is it.
 D431: That is not too bad.
 D432: Check over and see what we have on the list,
 D433: and see if they have red cabbage.
 D434: That is it on the list.
 D435: OK, that is about it for today.

Once the protocol has been broken up into small phrases, a major issue in protocol analysis is whether the verbalizations should be further encoded into formal categories. Newell and Simon (1972) argue that encoding the verbalizations can result in a substantial loss of information since language is already a device for encoding information. Consequently, they emphasize dealing with a protocol in the form of small phrases. Although treating protocols at this level of analysis can prove valuable in consumer research (e.g., Payne, 1976), a detailed and formal analysis of verbal data will probably require some form of encoding the verbalizations.

In the present study, each phrase in a protocol was coded in terms of a product category code and a process code.² There were 41 separate product category codes, generated from lists of product categories found in the supermarket trade journal *Progressive Grocer*. Because not all statements referred to a particular product category, three additional codes were used to reflect shopper movement statements, general comments about supermarkets, and miscellaneous statements.

The statements were also assigned one of 20 process codes grouped under five general headings: (1) statements of goals, needs, or strategies, e.g., "Do I need some butter"; (2) statements of simple awareness, such as item identification or item location, e.g., "There's the 2% milk"; (3) statements concerning information processing within product categories (levels of product attributes, absolute evaluations of product attributes or products, comparison of products), e.g., "This is 3 for \$99." "This is too sweet." Coke tastes better than Pepsi."; (4) product choice decisions, e.g., "OK, I'll buy the M&M's with peanuts."; and (5) miscellaneous statements.

The coding scheme used was based on two considerations. First, any coding strategy must reflect the nature of the task environment. In our case, coding the statements in terms of product categories emphasizes the organizing principles of all supermarket environments. Second, a scheme for coding must also reflect the theoretical interests and concerns of the researchers.

²The results to be presented are based on a partial analysis of our data base-35 protocols reflecting the behavior of 10 shoppers. The protocol analyses are based on approximately 20,000 coded statements.

For example, two important issues affecting the coding of statements were the amount of absolute evaluation versus comparative judgment and the amount of evaluation involving self and involving others. Absolute evaluations against internal or external standards and direct comparison of alternatives on their respective attributes has been identified as a crucial distinction in decision theory (Pitz, 1975; Payne and Braunstein, Note 4). The potential importance of viewing the shopper as a family purchasing agent has been stressed by Davis (1976). Consequently, our coding scheme distinguished between statement referring to the shopper's belief and those referring to the beliefs of others. Other aspects of the coding scheme, including the emphasis on shopper strategies, were derived from similar considerations. The amount of attention given to different product categories was measured by the number of times an attribute was mentioned with reference to the product category. Across all shoppers, the product categories with the largest number of references included: fresh meat, fresh fruits and vegetables, dairy products, paper products, canned fruits and vegetables, cookies and snacks. The categories with the fewest mentions included canned meats, pet foods and supplies, specialty foods, and candy, gum, and nuts. A ranking of product categories by the number of attribute references proved to be similar to the ranking of product categories by the percentage of shoppers purchasing in each category (see *Progressive Grocer*, January, 1964; October, 1975).

Table 1 presents a breakdown of the product attributes mentioned by the shoppers. Notice that the most frequently mentioned product attribute was price. This was true for all shoppers and for all but three product categories (specialty foods, breakfast cereals, and desserts). This result supports earlier studies that indicate price sensitivity among shoppers (e.g., Russo, Krieser, and Miyashita, 1975). However, the result conflicts with studies that have found shoppers to be poor estimators of product prices (see *Progressive Grocer*, 1974). One possible explanation is that shoppers do not store all price information in memory (where it can be retrieved when asked for prices in an interview before or after shopping). Instead, shoppers make use of the store environment (with prices displayed on every package) as form of external memory.

TABLE 1
 PRODUCT ATTRIBUTES MENTIONED DURING GROCERY SHOPPING

Product Attributes	Percentage of Mentions
Price (specific mention of price, item on sale, special, coupon)	45.7%
Size, Quantity	16.1
Form, Type	7.5
Brand	5.9
Ingredients	4.7
Taste, Flavor	4.7
Freshness, Appearance	3.2
Color	2.5
Package	2.0
Other	7.7
Number of Mentions of Product Attributes ¹	4,160

1. This analysis is based upon mentions of product attribute levels only; absolute evaluations and comparisons of product attributes were excluded.

This explanation is only speculative, but it does point up one of the possible advantages of in-store protocols as opposed to shopper interviews.

Table 2 presents a summary of the types of process statements. The results appeared to be reasonably stable across individual shoppers. Measured in terms of amount of verbal activity, shoppers made relatively little effort comparing brands within product categories while shopping (2.1% of all statements). One possible interpretation of this result is that while in the store, shoppers may decide whether or not to purchase in a particular product category (e.g., to buy rice for dinner), but once the decision is made, the brand choice is more or less routine. A related interpretation is that a substantial amount of shopper behavior in super-markets may involve processing at the product choice level in a shopper's goal hierarchy, and only a limited amount of processing may occur at the brand choice level.

The protocols also provided evidence supporting the concept that housewives seem to act as purchasing agents for their families. For example, Table 3 present excerpts from the protocols relating candy buying to other family members.

Additional information on shopper behavior was obtained through direct observation and purchase diaries. Table 4, for instance, presents the relationships between the use of a shopping list and the number of items and amount purchased. The data presented in Table 4 are based on the entire sample of 19 shoppers. Notice that when lists were used, only approximately 41% of the items purchased were found on the list. The percent of items on the list to total items bought per trip ranged

TABLE 3

EXAMPLES OF SHOPPER COMMENTS RELATING CANDY BUYING TO OTHER FAMILY MEMBERS

"It's very difficult to buy candy at this time of year that satisfies my husband's sugar tooth, because he likes chocolates and they damage and then he cannot hold them. They melt. He loves M&M's."

"Everybody likes these caramel creams. And what does my husband like? He eats one candy bar--Butterfingers. These are awful."

"I do not think my husband will chew the sugarless stuff. He likes Juicy Fruit which I cannot stand; but I will buy it for him."

"Barnum's Animals, Chuckles, Jujubees. This is my husband's favorite kind of candy; it pulls fillings out of teeth, but it's really good while going down--you know, jelly. I think I'll buy him some animals. He'll be trilled by that, I know."

(Mother shopping with child)
 Mother: "Do you like the peanuts?"
 Child: "I like the plain."
 Mother: "The plain ones. All right, we will get the plain M&M's."

TABLE 2

NATURE OF VERBAL ACTIVITY DURING SHOPPING

--by individual shopper--

Nature of Statement	Shopper										
	1	2	3	4	5	6	7	8	9	10	\bar{X}
<u>Within product category statements</u>	33%	34%	33%	28%	33%	34%	28%	40%	41%	32%	33.4%
Product Attributes	(24)	(20)	(19)	(17)	(22)	(21)	(16)	(24)	(25)	(23)	21.0%
Absolute Evaluations	(8)	(11)	(12)	(10)	(8)	(12)	(10)	(14)	(12)	(7)	10.3%
Brand Comparisons	(1)	(3)	(2)	(1)	(3)	(1)	(2)	(2)	(4)	(2)	2.1%
<u>Awareness</u>	29	18	12	22	20	16	21	20	13	19	19.8%
<u>Need or strategy statements</u>	13	16	18	18	13	20	19	18	13	16	16.1%
<u>Explicit buy or not-buy statements</u>	4	8	8	7	5	9	8	8	10	4	6.9%
<u>Miscellaneous statements</u>	21	24	29	25	29	21	24	14	23	29	23.8%
Number of Statements	2965	2129	1967	2287	2389	1821	1581	1928	1433	1303	19,803

TABLE 4

RELATIONSHIP BETWEEN USE OF SHOPPING LIST
AND AMOUNT PURCHASED

--by individual shopper--

Shopper	Number of Trips ¹		Average \$ Purchase per Trip		Average No. of Items Purchased per Trip		% of Items on List To Total Item Purchased
	With List	Without List	With List	Without List	With List	Without List	
1	6	0	\$37.84		27		76%
2	3	3	44.27	40.27	46	35	50
3	5	0	46.21		43		54
4	5	0	25.94		25		49
5	1	4	41.55	44.01	39	41	26
6	4	2	35.05	40.27	34	32	29
7	3	2	60.25	52.32	53	49	
8	0	5		75.13		64	
9	3	3	48.47	15.58	43	17	22
10	0	3		38.85		42	
11	3	0	28.42		21		70
12	5	0	42.86		29		39
13	2	4	97.70	88.83	75	70	32
14	3	3	56.47	50.71	33	29	29
15	2	3	52.28	44.87	48	46	27
16	4	2	43.66	45.06	47	50	38
17	0	1		33.34		31	
18	0	5		57.75		42	
19	3	2	76.39	41.52	49	34	29
Average ²			45.78	51.39	38	44	41
Average ³			54.07	48.00	46	41	

1. Trips with incomplete data were excluded from these analyses.
2. All shoppers were included in this analysis.
3. Shoppers who purchased only with or without lists were excluded from this analysis.

from 22% to 76% across shoppers.

The extent of buying from end-of-aisle and free standing displays was also determined by combining data from direct observation and the record of purchases. It was found, for example, that taking all shoppers together, 14% of all purchases were from displays. Looking at individual shoppers, the proportion of items purchased from displays ranged from 1% to 25%. However, these results should be interpreted with caution. The results may reflect differences among shoppers in terms of strategies, or could be due to differences among the stores' use of point-of-purchase materials.

Summary

The results presented here illustrate how verbal protocols, in conjunction with other, more traditional marketing research tools, provided information about consumer decision behavior at a level of detail difficult to achieve otherwise. Data was obtained on the goals and attributes used by supermarket shoppers, e.g., price awareness; the choice processes used by shoppers, e.g.,

amount of brand comparison; and the effect of in-store and social influences on shopping behavior. The collection of such data, along with other forms of high density, sequential data seems essential if we are going to develop and test complex and dynamic models of consumer behavior.

With regard to other process tracing techniques, there are differences of opinion as to whether such methods as eye movement recording or explicit information search procedures, should be preferred over verbal protocols as a method of collecting process data. Russo and Rosen (1975) argue that eye-movement records are better because "They are unobtrusive, detailed, and difficult to misrepresent." These claims may be true, although it is not clear how unobtrusive eye-movement recording has been given the apparatus used. Simon (1976), on the other hand, has argued that "verbal protocols provide a rich source of data at densities comparable to the densities of eye-movement records, and containing far more information per data than the latter" (p.261).

Jacoby (1976) has argued even more strongly for the use

of information search techniques as opposed to verbal protocol procedures. He apparently feels that information search data are "behavioral" whereas verbal protocol data are not to be trusted. Jacoby does not seem to make any distinction between verbal protocols and other types of verbal data, such as responses to post-decision questionnaires and interviews. However, as was noted earlier, there are important distinctions between verbal protocol data and post-decision verbal reports. These distinctions make verbal protocols a much more valid source of process information.

One problem with the information acquisition methods is that they focus exclusively on the subject's use of objective, external information. The method does not easily allow for insights into an individual's use of information stored in internal memory. This problem relates to Simon's (1976) statement that verbal protocols provide more information about cognitive processes than eye-movement records.

Another problem is that information acquisition methods provide little or no indication of when and if the information being acquired is actually being processed. Payne and Braunstein (Note 4), for example, found one subject who (according to a monitoring of the subject's acquisition of information) appeared to be consistently searching all available information in an interdimensional fashion. The search appeared to be independent of all number of alternative gambles presented. Such a pattern of information use suggested that some form of an additive decision rule was being employed by the subject (see Payne, 1976). However, examination of the protocol showed that some of the acquired information was ignored by the subject and that the amount of information (number of attributes or dimensions) considered per alternative varied across alternatives; this pattern suggested a conjunctive type of model. Excerpts from the protocol made it clear that the subject had adopted the practice of obtaining all the available information before proceeding with the decision process. The ability of verbal protocols to detect such activity represents one of the greatest advantages of protocol over eye-movement and explicit information acquisition procedures.

Finally, information acquisition studies usually require the decision task to be more structured. This may present problems as consumer researchers attempt to do more "real-world" research.

Obviously, the preferred research strategy would be to employ more than one process tracing technique. Examples of such a multimethod approach to research can be found in Payne and Braunstein (Note 4) and Russo and Doshier (Note 2). Where a multimethod approach is impossible, the researcher will have to choose among process tracing methods on the basis of theoretical concerns, the availability of apparatus, and the relative difficulty in analyzing data collected by each method. The researcher faced with this choice can find encouragement in the results of efforts to compare methods. Newell and Simon (1972), for example, present evidence of a high degree of convergence in the interpretations about problem solving based on eye-movement data and on verbal protocols.

Finally, while we have argued the values of verbal protocols, it is clear that there are definite costs associated with the use of such a process tracing method. Protocols for example, provide researchers with a large amount of detailed data, which at present cannot be simply run through a computer program for analysis. While procedures for examining protocol data are currently being developed, it is clear that an investigator choosing to do a protocol study will probably have to settle for an intensive investigation of a

relatively small number of subjects. However, as illustrated by the efforts to investigate contingent processing in decision making (Payne, 1976; Payne & Braunstein, Note 4), it does appear that process tracing techniques can lead to generalities about behavior, in spite of the limited number of subjects run in any particular experiment.

Another problem with using verbal protocols is that standard summary statistics for such data are not well developed. Consequently, a researcher using such methods is often forced to present the results of his study in great detail, e.g., parts of the protocols themselves. Haines (1974), in fact, advocates such a procedure. Finally, it is obvious that verbal protocols represent an obtrusive measure of behavior. We have examined the obtrusiveness of the verbal protocol procedure as part of two studies of decision making (Carroll & Payne, 1977; Payne & Braunstein, Note 4). Our results seem to indicate that the verbal protocol procedure may slow down the decision process slightly but does not change it fundamentally. A similar effect for verbal protocols has been suggested by other researchers (e.g., Dansereau and Gregg, 1966; Newell and Simon, 1972). However, much more research is needed to examine directly the effects of various process tracing data collection methods on the behavior exhibited by subjects.

In summary then, verbal protocols are a valuable source of data for consumer research. However, it does not appear that by themselves they will "Save The World."

Reference Notes

1. Svenson, O. Coded think-aloud protocols obtained when making a choice to purchase one of seven hypothetically offered houses: Some Examples. Progress report on project of cognitive processing in decision making. University of Stockholm, 1974.
2. Russo, J.E., Doshier, B.A., Dimensional evaluation: A heuristic for binary choice. Unpublished manuscript, University of California, San Diego, 1975.
3. Payne, J.W., Braunstein, M.L., & Carroll, J.S., Exploring predecisional behavior: An alternative approach to decision research. Unpublished manuscript Graduate School of Business, University of Chicago, 1977.
4. Payne, J.W., & Braunstein, M.L. Contingent processing in risky choice: A process tracing investigation. Unpublished manuscript, Graduate School of Business, University of Chicago, 1977.

References

- Bettman, J.R. Information processing models of consumer behavior. Journal of Marketing Research, 1970, 7, 370-376.
- Bettman, J.R. Decision-net models of buyer information processing and choice: Findings, problems, prospects. In G.D. Hughes and M.D. Ray (Eds.), Buyer/consumer information processing. Chapel Hill, North Carolina: University of North Carolina Press, 1974.
- Carroll, J.S., & Payne, J.W. Judgments about crime and the criminal: A model and a method for investigating parole decisions. In B.D. Sales (Ed.), Perspectives in law and psychology. (Vol. 1): The criminal justice system. New York: Plenum, 1977.

Dansereau, D.F., & Gregg, L.W. An information processing analysis of mental multiplication. Psychonomic Science, 1966, 6, 71-72.

Davis, H.D. Decision making within the household. Journal of Consumer Research, 1976, 2, 241-260.

Haines, G.H. Process models of consumer decision making. In G.D. Hughes & M.L. Ray (Eds.) Buyer/consumer information processing. Chapel Hill, North Carolina: University of North Carolina Press, 1974.

Jacoby, J. The emerging behavioral process technology in consumer decision making research. In W.D. Perreault, Jr. (Ed.), Advances in consumer research (Vol. 4). 1977, 263-265.

King, R.H. A study of two problems of building a model to simulate the cognitive processes of a shopper in a supermarket. In G.H. Haines (Ed.), Consumer behavior: Learning models of purchasing. New York: The Free Press, 1969.

Montgomery, H.A. A study of intransitive preferences using a think aloud procedure. In H. Jungerman, & G. de Zeeuw (Eds.), Proceedings of The Fifth Research Conference on Subjective Probability, Utility and Decision Making, 1976.

Newell, A., & Simon, H.A. Human problem solving. Englewood Cliffs, New Jersey: Prentice-Hall, 1972.

Nisbett, R.E., & Wilson, T.D. Telling more than we can know: Verbal reports on mental processes. Psychological Review, 1977, 84, 231-259.

Payne, J.W. Task complexity and contingent processing in decision making: An information search and protocol analysis. Organizational Behavior and Human Performance, 1976, 16, 366-387.

Pitz, G.F. Decision making and cognition. In H. Jungerman & G. de Zeeuw (Eds.), Proceedings of The Fifth Conference on Subjective Probability, Utility, and Decision Making, 1976.

Russo, J. E. Krieser, G. & Miyashita, S. An effective display of unit price information. Journal of Marketing, 1975, 39, 11-19.

Russo, J.E., & Rosen, L.D. An eye fixation analysis of multialternative choice. Memory & Cognition, 1975, 3, 267-276.

Simon, H.A. Discussion: Cognition and social behavior. In J.S. Carroll & J. W. Payne (Eds.) Cognition and Social Behavior, Hillsdale, New Jersey: Erlbaum, 1976.

Tversky, A. Intransitivity of preferences. Psychological Review, 1969, 76, 31-48.

THE USE OF THE ACTIVE PROFILE EVALUATION PARADIGM IN
STUDYING CONSUMER JUDGMENT PROCESSES

Peter Wright, Stanford University
Peter Rip (student), Stanford University

A number of methods have recently been developed in the study of consumer information processing. In this paper we review how the Active Profile Evaluation paradigm has been used as one method of inquiry. We will also attempt to outline alternative applications of the paradigm in the study of consumers' judgment processes.

The paradigm requires the decision maker (DM) to rate or rank order several real or hypothetical alternatives. The profiles are generally composed of few (2-6) cues, with product dimensions generally referring to tangible characteristics. The analysis then proceeds to relate variations in a consumer's preferences to variations in the product characteristics. A variety of estimation procedures are available; all yielding an index of goodness-or poorness-of-fit and estimates of the relative value the DM assigned to the various attribute levels. Within this paradigm, researchers have asked a number of questions. The most popular issue has been the comparison of relative fits of various mathematical representations of evaluative strategies. The criteria of fit have generally been the multiple correlation coefficients, stress values, and cross-validation coefficients. Since this is an individual-level method of analysis, one generally examines these statistics using cross-sectional designs. One such approach would be to ask whether systematic individual differences, situational, or psychological factors (e.g. risk, distraction) lead to better fits for certain non-linear integration processes. Examples of this type of analysis are papers by Einhorn (1971) and Wright (1974; Scott and Wright, 1976).

The general linear model (OLS, ANOVA) has been the predominant technology associated with this paradigm (Slovic and Lichtenstein, 1971). Recent studies in marketing have also used non-metric estimation methods (e.g. linear programming [Parker and Srinivasan, 1976; Pekelman and Sen, 1974] and conjoint measurement [Green and Wind, 1973]). One consequence of these technologies has been the remarkable robustness of linear additive models of judgment, naive theories notwithstanding. Much of the early research focused upon the detection of non-linearities in judgment processes (Einhorn, 1971; Slovic, 1969; Slovic, Fleissner, and Bauman, 1971). The apparent robustness of the arithmetically simple, but cognitively complex, linear additive models now seems to be more an artifact of the model, rather than any real congruence with mental processes (Birnbaum, 1973; Einhorn and Hogarth, 1974; Wainer, 1975).

The search for appropriate statistical representations of choice processes naturally invites comparisons of alternative models of the same behavior. A number of formal approximations of alternative choice strategies have been developed (e.g. Einhorn [1971]). Typically, the approach has been to estimate the parameters of a variety of different models and select the one with the lowest mean squared error or stress as the best representation. This criterion of selection has been severely criticized (Birnbaum, 1973) because it artificially favors certain model forms.

Aside from the mere inspection of multiple correlations or cross validation coefficients of alternative models, two related validation methodologies have developed in the search for the "appropriate" paramorphic representation (Hoffman, 1960). Neither approach has been applied

extensively in the consumer behavior literature, for reasons which will be discussed later. One method is based upon the axiomatic analysis of choice rules (Krantz and Tversky, 1971; Luce and Tukey, 1964). This method reduces a choice rule to a set of unique axioms, which can then be tested by inspection of the data. The analysis is diagnostic, rather than statistical, searching for alternative models which best account for the empirical results. The principal limitation of this method of assessment is the lack of an error theory (Krantz and Tversky, 1971) which allows us to distinguish minor perturbations in the data from systematic deviations. An alternative, but related, method of validation is Anderson's (1974) functional measurement technology. The details of the method are available from a variety of sources (Anderson, 1974; Slovic and Lichtenstein, 1971). The method, essentially, achieves a monotonic rescaling of the dependent variable such that the postulated model can usually be tested in an ANOVA framework. Although Anderson and colleagues have applied the technology across a number of tasks, there have been few applications of functional measurement in the consumer behavior literature. Exceptions are the work of Bettman, Capon, and Lutz (1974; 1975) and Troutman and Shanteau (1976).

Aside from the question of model comparisons, the weight and utility estimates derived from the procedure can be used as indirect measures of those factors which drive behavior. In a more familiar sense, we can either ask consumers for their attitudes directly or supplement their self-reports with indirect or derived estimates of their attitudes. These indirect methods of assessment may often be more reliable than more obtrusive techniques. Application of the active profile evaluation paradigm in these situations has been less common, but is potentially more powerful. Explicitly, if one's research pertains to certain aspects of consumers' product evaluation strategies, the weights and utilities are directly relevant. Using estimated weights or utilities as dependent variables, rather than global fits, it becomes less critical for these weights to be derived from a model which is a "true" description of the process. The research emphasis shifts from absolute fit to one of between- or within-subject differences in weights or utilities. Applied in this way, the active profile evaluation paradigm is simply a measurement technique for tapping cognitive events. Our concern for the reliability and validity of the measures obtained should parallel the concern we show over these issues with respect to other techniques.

One of the most interesting issues concerns a person's self-insight with respect to his or her own judgment processes. Evidence from a variety of sources (Cook and Stewart, 1975; Nisbett and Wilson, 1977; Scott and Wright, 1976; Slovic, Fleissner, and Bauman, 1972) suggests that subjects tend to report their behavior was influenced by more than the small subset actually observed. In cases of disagreements between derived parameters and subjective reports, one looks for systematic patterns in the deviations. The more thorough one's tests of reliability and validity of model-fitting analysis, the more confident one feels in attributing systematic differences to reporting biases. Observed discrepancies between self-reports and derived parameters have inspired further analyses of the source of the

reporting bias and the meaning of "relative importance" to naive judges (Cook and Stewart, 1975; Scott and Wright, 1976).

Until now we have discussed self-insight with respect to relative weights among dimensions. However, we may use estimated utilities within a dimension to assess how well consumers can recall preference orderings along that dimension. Interrogating consumers about relative preferences as just stated, rather than introspective reconstruction of cognitive processes, may prove to be a very fruitful area of research. This allows us to directly examine the extent of various cutoff-types of strategies in consumer choice. For example, we might expect that someone with accurate self-insight about his or her use of absolute cutoffs would demonstrate certain non-linearities within that dimension.

Often the research question may concern the effects of specific treatments upon generic aspects of the product evaluation strategy. Examples would be its complexity or degree of unidimensionality, or the presence of certain risk preferences or aversions. The first question may be addressed by comparing the relative weights that consumers gave to their two most important dimensions (Wright and Weitz, 1977), or the variance across the set of weights. The second question may be addressed by comparing gaps between the utilities of specific levels of a dimension. For example, one recent study found that women showed much larger decreases in utilities as the chance of negative outcomes increased modestly. This effect was particularly strong for those consumers for whom the outcome seemed imminent relative to when it seemed more distant.

A related question of interest to consumer researchers is how DMs account for others' preferences when making their own choices. Clearly, self-insight and awareness of others' utility functions are relevant to this point.

The actor-observer distinction (Jones and Nisbett, 1971; Ross, 1977) could serve as a useful model in studying how we assess others' utilities. The study of how we judge and integrate others' preferences is one area of basic consumer research which may be potentially fruitful.

A question related to the general problem of assessing others' utility functions is the degree to which group judgments reflect preferences of the individual members. One approach to this question is to examine the correspondence between the individual members' utilities and those of the group outcome as derived from a group rating or ranking of hypothetical alternatives. While this method does not offer direct evidence concerning group decision making processes, it does offer a means of examining members' impact in terms of measures of net effect, rather than process-oriented criteria such as the number of influence attempts. Indeed, one of the major advantages of the active evaluation paradigm is its ability to complement process-oriented measures (e.g. protocols, eye movements) with rigorous behavioral measures of outcomes.

A final highly promising application of the active profile evaluation paradigm is in assessing the effects of alternative persuasive messages or other types of information displays upon consumers' evaluation strategies. If alternative messages and information displays impact upon such things as cue weights, relative or absolute attention, cutoff usage, attribute utilities, complexity of choice strategy, and accuracy of interpersonal assessment of utilities, the active evaluation of profiles promises to be one useful way of tracking these effects.

While the active profile evaluation paradigm is quite flexible and a potentially attractive measurement tech-

nique, one limitation seems noteworthy. The active evaluation paradigm requires multiple judgments in order to derive reliable parameter estimates. Often the number of judgments can be reduced by the use of fractional designs. However, ten to thirty judgments may still be necessary for reliable parameterization. Therefore, there may be a strong task effect; biasing subjects toward simpler rules and candidate-wise processing. So the evidence produced may not generalize too well beyond situations in which people make a number of judgments in a limited period of time. A researcher can control the time span over which the succession of judgments is made, thereby replicating a setting in which rapid or highly dispersed judgments are made. Perhaps our ingenuity and flexible use of this task factor will allow us to minimize the potential for bias introduced by this essential task factor.

In conclusion, the versatility of the active profile evaluation paradigm should not be ignored. We continue to recommend that a hypothesis worth testing is worth testing in several ways, and the outputs from this paradigm should be viewed from this perspective. Furthermore, the paradigm provides important confirmatory behavioral evidence in a task which is fully compatible with other consumer choice technologies.

References

- Anderson, Norman H., Information Integration Theory: A Brief Survey, in D. H. Krantz, R. C. Atkinson, R. D. Luce, and P. Suppes (eds.), Contemporary Developments in Mathematical Psychology, Vol. II, San Francisco; W. H. Freeman, 1974.
- Bettman, James R., Noel Capon, and Richard J. Lutz, Cognitive algebra in multi-attribute models, Journal of Marketing Research, 12, (May 1975), 151-164.
- Bettman, James R., Noel Capon, and Richard J. Lutz, Multi-attribute measurement models and multiattribute attitude theory; a test of construct validity, Journal of Consumer Research, 1, (March 1975), 1-15.
- Birnbaum, Michael H., The Devil Rides Again: correlation as an index of fit, Psychological Bulletin, 79, 1973, 239-242.
- Cook, Richard C., and Thomas R. Stewart, A comparison of seven methods for obtaining subjective descriptions of judgment policy, Organizational Behavior and Human Performance, 13, (February, 1975), 31-45.
- Einhorn, Hillel J., Use of nonlinear, non compensatory models as a function of task and amount of information, Organizational Behavior and Human Performance, 6, (January, 1971), 1-22.
- Einhorn, Hillel J., and Robin M. Hogarth, Unit Weighting Schemes for Decision Making, Organizational Behavior and Human Performance, 13, 1975, 171-192.
- Green, Paul and Yo'ram Wind, Multiattribute Decision Making in Marketing, Englewood Cliffs: Prentice Hall, 1973.
- Hoffman, Paul J., The Paramorphic Representation of Clinical Judgment, Psychological Bulletin, 57, 1960, 116-131.
- Jones, E. E. and R. Nisbett, The Actor and the Observer: Divergent Perception of the Causes of Behavior, in E. E. Jones et al (eds.), Attribution: Perceiving the Causes of Behavior, Morristown, N.J.; General Learning Press, 1971.
- Krantz, David and Amos Tversky, Conjoint-Measurement Analysis of Composition Rules in Psychology, Psychological Review, 78, 1971, 151-169.

Luce, R. D. and J. W. Tukey, Simultaneous Conjoint Measurement; a new type of fundamental measurement, Journal of Mathematical Psychology, 1, 1964, 1-27.

Nisbett, R. and T. D. Wilson, Telling More Than We Know: verbal reports on mental processes, Psychological Review, 84, 1977, 231-259.

Parker, Barnett R. and V. Srinivasan, A Consumer Preference Approach to the Planning of Rural Primary Health-Care Facilities, Operations Research, 24, 1976, 991-1025.

Pekelman, Don and Subrata K. Sen, Mathematical Programming Models for the Determination of Attribute Weights, Management Science, 20, (April, 1974), 1217-1229.

Ross, Lee, The Intuitive Psychologist and his Shortcomings: Distinctions in the Attribution Process, in Berkowitz, L. (ed.) Advances in Experimental Social Psychology, Vol. 10, 1977, New York; Academic Press.

Scott, Jerome E. and Peter Wright, Modeling an Organizational Buyer's Product Evaluation Strategy; Validity and Procedural Considerations, Journal of Marketing Research, 13, (August, 1976), 211-224.

Slovic, Paul, Analyzing the Expert Judge; A Descriptive Study of a Stockbroker's Decision Processes, Journal of Applied Psychology, 53, (August, 1969), 255-263.

Slovic, Paul, Don Fleissner, and W. Scott Bauman, Analyzing the use of information in investment decision making: A methodological proposal, Journal of Business, 45, (April, 1972), 283-301.

Slovic, Paul and Sarah Lichtenstein, Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment, Organizational Behavior and Human Performance, 6, (November, 1971), 649-744.

Troutman, C. Michael and James Shanteau, Do Consumers Evaluate Products by Adding or Averaging Attribute Information?, Journal of Consumer Research, 3, 1976, 101-106.

Wainer, Howard, Estimating Coefficients in Linear Models; It Don't Make No Never Mind, Psychological Bulletin, 83, (March, 1976), 213-217.

Wright, Peter, The Harassed Decision Maker: Time Pressures, Distractions, and the Use of Evidence, Journal of Applied Psychology, 59, (November, 1971), 555-561.

Wright, Peter and Barton Weitz, Time Horizons and Product Evaluation Strategies, Journal of Marketing Research, 1977, in press.

CHRONOMETRIC ANALYSIS: AN INTRODUCTION AND AN APPLICATION TO
LOW INVOLVEMENT PERCEPTION OF ADVERTISEMENTS

Meryl P. Gardner (student), Carnegie-Mellon University
Andrew A. Mitchell, Carnegie-Mellon University
J. Edward Russo, University of Chicago

Abstract

The methodology involved in using response times to examine internal information processing is briefly reviewed and the application of this approach to consumer research is discussed. A study is described that uses response times to examine the differential effects of high and low involvement learning of advertising messages. Low involvement learning is hypothesized to take place under either attention limited or strategy limited conditions. The experiment examined the latter condition where consumers fully attend to the advertisement but not for the purpose of evaluating the advertised good. The results of the experiment confirm several predicted effects of this type of low involvement processing, including the formation of more positive attitudes toward the advertised brand.

Introduction

All consumer decisions rely, to some extent, on information stored in memory. Recent research on consumer decision making, however, has deemphasized the role of memory in consumer decision processes by concentrating attention on external information acquisition. Since acquisition responses, such as reaching or looking, are conveniently observable, they offer a clear methodological advantage. In contrast, retrieval of information from memory is covert.

It seems clear, however, that if we are to understand consumer decision processes the important role of memory and internal processing cannot be excluded from experimental investigation. Currently, a consensus seems to be forming that the time is appropriate for consumer research to direct its attention at the internal processes of information encoding and storage and to their effect on consumer decision making (Bettman, 1978; Chestnut and Jacoby, 1977; Johnson and Russo, 1977; Olson, 1978).

Available Methodologies

The investigation of internal processes is hindered by the lack of associated overt behavior. There are, however, two observations that can always be made on a cognitive process: the output of that process and the time taken to complete it. Both of these measures provide a data base for methodologies that can successfully investigate internal processes. The focal point of this paper is on chronometric analysis or the use of response times to examine internal processes; however, it might be helpful to first discuss methodologies based on output. To use output effectively, the experimental environment has to be designed so that the observed response can discriminate among possible processes. This is called input-output analysis, or, in its most refined form, the axiomatic method (Chase, 1977).

The output of a process can be elaborated by two process tracing techniques. The more common requires subjects to generate a verbal report of their own cognitive processes. Technically, these verbal protocols are the output of a second process, that of self-commentary. Although often very informative, they are difficult to analyze formally and are susceptible to distortion when based on memory

rather than emitted concurrently with the primary task. Furthermore, verbal protocols are almost useless for brief or highly automatized cognitive processes of which subjects are not consciously aware (Nesbitt and Wilson, 1977). Unfortunately, much consumer behavior, such as encoding information from advertisements, may be of this nature.

A second elaboration of simple output analysis is the recording of intermediate responses, such as information acquisition behavior. This is still output analysis, except now the output of subprocesses is analyzed alongside that of the primary process. Recent consumer research has made excellent use of this technique. Through its application we have learned that people often ignore large amounts of product information, they adapt to the presentation mode or format of the information and they are flexible in their selection of decision strategies, seldom using a single, "pure" strategy. This progress has been achieved, however, in laboratory settings characterized by external sources of product information, and, when investigating known alternatives, has failed to take into account previously stored information (Bettman, 1978).

Apart from these two process tracing elaborations, the input-output framework includes a wide variety of specific experimental paradigms for the investigation of memory. Tulving and Bower (1974) provide a list of ten such techniques commonly used in cognitive psychology. These include: feature probing, recall intrusions, retrieval cueing, and clustering in recall. Although we shall now turn out attention to chronometric analysis, the power and diversity of input-output analyses should not be ignored.

Chronometric Analysis

As mentioned previously, the purpose of chronometric analysis is to use response times to examine internal information processes. In this context, response times (RT's) are generally defined as the amount of time between the presentation of a stimulus to an individual and his or her response to that stimulus.

Chronometric analysis has been used extensively by cognitive psychologists to examine a wide range of mental processes. It has been used, for instance, to examine the underlying mental processes involved in verifying that a presented stimulus belongs to a previously memorized set (Chase and Clark, 1972; Clark and Chase, 1972) and the quantification of visual stimuli (Klahr and Williams, 1976).

The fundamental principle of chronometric analysis is that response time measures the amount of processing that has taken place. In order to do this we generally have to assume that the processes are executed in series rather than in parallel. For many applications this principle alone is sufficient. Subjects, for instance, may be asked to respond to the same stimulus under different conditions that imply different response times. The observed RT's are then used to test the theory from which the predictions were derived.

Before we discuss general procedures for using chronometric analysis to examine internal processes, it will be helpful to present a specific example of its use from the cognitive psychological literature. Collins and Quillian (1969) used response times to test Quillian's (1969) model of semantic memory. According to the model, semantic memory can be represented as a hierarchical network with each node representing an object or concept. An example of such a hypothetical semantic network¹ for cola and lemon lime soft drinks is presented in Figure 1. Moving

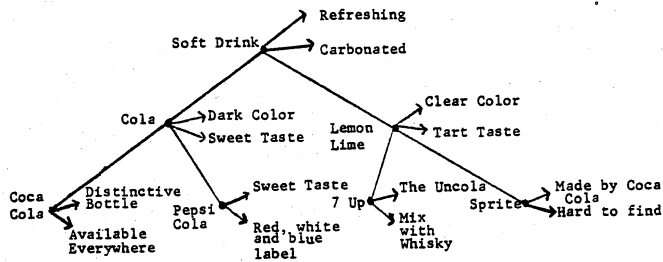


FIGURE 1. HYPOTHETICAL THREE LEVEL HIERARCHY FOR COLA AND LEMON LIME SOFT DRINK

to a higher level in the network signals a superset relation. For example, cola is the superset of Pepsi Cola, and soft drink is the superset of cola.

The power of this conceptualization is both the efficiency of information storage, since information is stored at its highest level of generality and the ability to infer or verify novel facts. For example, an individual does not have to store both "Coca Cola has a dark color" and "Pepsi Cola has a dark color"; this information is stored with the concept cola. Also, the sentence "Pepsi Cola has a dark color" can be verified even though that particular association has never been encountered.

According to the model, the verification process requires a search through the nodes of the semantic network. For instance, verifying "Coca Cola has a stimulating taste" would require activating the links from Coca Cola to cola to soft drink where the property "has a stimulating taste" can be accessed.

To test the model, RT's were recorded as subjects verified statements that involved searching different levels of the hierarchy. In general, the results supported the model.² The more hierarchies that needed to be searched to respond to a question, the longer the response time. (Other models of memory have since been shown to be compatible with the results; see Chase, 1977³). The response times also provided estimates of the time needed to traverse between nodes (about 75 msec) and to retrieve a property at a node (about 225 msec). The former can be measured by subtracting the times to verify sentences

¹In Collins and Quillian's experiment, the structure of semantic memory for animals was examined by using the subsets of birds and fish.

²The data for positive responses fit the model well, however, this was not true for negative responses. It was hypothesized that different strategies may be involved in falsifying sentences.

³For instance, distance in semantic network is confounded with past associative frequency and, therefore, with associative strength.

like "Sprite is lemon lime" from "Sprite is a soft drink" or "Sprite is made by Coca Cola" from "Sprite has a clear color." Each pair of sentences involves the same processing except for an additional level change in the second sentence.

The Subtractive Method

The Collins and Quillian experiment illustrates the use of the subtractive method which was first proposed by Donders (1868). Logically the method is quite simple. If a number of mental processes occur serially, then by eliminating one or more of the processes in a particular experimental condition, one may obtain a measure of the time required for the remaining processes. In the Collins and Quillian experiment, the serial operations are an original retrieval of a concept from memory and then a sequential search through the different hierarchies. The occurrence of this relationship in an actual experiment provides evidence supporting the hypothesized model.

The validity of the subtractive method relies on an assumption that might be called the Principle of Process Deletion. It is assumed that the process of interest can be deleted from some sequence of processes without affecting any of the others. That is, the deleted process does not interact with any other process. (Technically, the durations among processes must not interact.) This assumption is critical to the validity of the subtractive method but, unfortunately, is not always testable.

From the above discussion, it should be clear that the subtractive method requires an adequate theory of the process that is to be examined. In order to construct a task that deletes only the specific process of interest, one should know the sequence of cognitive activities that occurs between stimulus and response. At the least, one must know enough to be certain that the Principle of Process Deletion is applicable, but usually this knowledge is only part of a theory of the complete process.

The Additive Factor Method

The additive factor method, proposed by Sternberg (1969), is an elaboration of the subtractive method. Its principal concern is to provide procedures for developing a conceptual model of the cognitive processes that intervene between stimulus presentation and the subjects' response. Whereas the subtraction method requires a well-developed theory in order to interpret the resulting reaction times, the additive factor method is designed to aid in the development of a theory of the intervening processes.

The basic logic underlying this method is that each reaction time interval contains a series of independent stages. Each stage of the process transforms the information received from the previous stage and outputs it to the next stage. It is assumed that the duration of each stage is independent of the stages that precede it. The implications of this conceptualization are as follows. First, the total reaction time of a particular process is the sum of the stage durations. Second, independent effects on total reaction time will be produced whenever experimental manipulations (independent variables) affect different stages. Finally, if the effects of two experimental manipulations interact in a statistical sense then they must affect the same stage. Thus, in using the additive factor method, a multifactor experiment is applied to a particular information processing task. Each pair of factors that have an additive effect on reaction time affect different stages of the process, and each pair that interact affect the same stage. Consequently, in order to obtain a precise analysis of the stages for definitional purposes, more than a few factors must be manipulated.

To illustrate the additive factor method, we rely on the experiments of Sternberg (1969) which investigated short term character recognition. In these experiments a subject was required to memorize a set of stimuli and then, upon presentation of a stimulus, to respond positively or negatively depending on whether it was contained in the previously memorized set. The factors manipulated in these experiments included: pictorial clarity of the presented stimulus, size of the previously memorized set, whether a positive or negative response was required, and the relative frequency of either a positive or negative response. The results of these experiments indicated that all four factors had a significant effect on reaction time and that all six pairwise interactions were either insignificant or there was a compelling a priori reason for not expecting a statistically significant interaction. Consequently, Sternberg (1969) hypothesized that the task involved four independent stages, each associated with one of the factors tested: encoding, comparison, response choice and response execution. Note that four stages are the most that could be discovered in a four factor design and that there could possibly be more stages to the process that could not be identified in this experiment.

Methodological Problems

As with any technique, chronometric analysis has accompanying methodological problems. The primary problem is the natural trade-off between speed and accuracy. Subjects can usually perform faster if they sacrifice accuracy or more accurately if they sacrifice speed. This relation is captured by the speed-accuracy operating characteristic (Pew, 1969). An idealized version of this function, adapted from Pachella (1974), is shown in Figure 2.

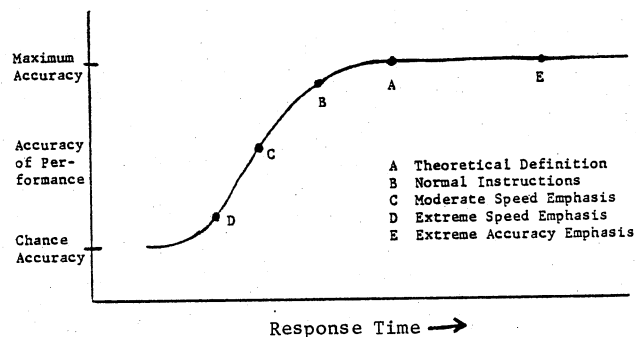


FIGURE 2. IDEALIZED SPEED-ACCURACY OPERATING CHARACTERISTIC AFTER PEW (1969)

According to the definition of reaction time, subjects should be responding at point A. However, they generally operate below this point. Most of the problems associated with the speed-accuracy trade-off can be avoided if subjects perform at the same accuracy level under all experimental conditions. Then the relative values of the observed RT's will reflect the effect of the independent variables and not the effect of a different accuracy strategy. When accuracy fails to remain constant across conditions, there is a danger that incorrect inferences will be drawn from the RT data. (For an example, see Pachella, 1974). A number of procedures have been proposed for adjusting reaction times if different accuracy levels occur in an experiment. A critique of these procedures may be found in Pachella (1974).

The second problem to be discussed concerns the interpretation of response times. Whenever detailed processes are manifested only in behavior as aggregate as response time, the theoretical conclusions must rely heavily on the theoretical input. In spite of the power of the ad-

ditive factors method, it still requires reasonable hypotheses about potential stages of the process in order to design an appropriate multifactor experiment.

Prior to Sternberg's introduction of the additive factors method, the subtractive method had fallen in disfavor. In discussing this particular period, Chase (1977) states, "In retrospect, it appears that the lack of success with the subtractive technique was due primarily to the absence of good cognitive theory." For consumer researchers interested in relatively complex tasks, this means that RT's are not the proverbial free lunch. Even the most sophisticated chronometric analysis can only refine theory, not create it.

Response Time and Consumer Behavior

There are a number of different areas in consumer behavior where the use of reaction times would seem to be useful. One area is the structure of product knowledge (i.e., the organization of memory for product information). A number of questions are of interest here. What is the structure linking brands and attributes within a product class? Is it hierarchical as pictured in Figure 1, or organized in accord with the more recent models of semantic memory (e.g., Anderson and Bower, 1973, or Norman and Rumelhart, 1975)? Does brand name play a central role linking all brand associations, as some researchers have argued? What is the relative efficacy of brand-based versus attribute-based memory organizations? When the structure of product knowledge is identified, will it be invariant across individuals and product categories? These are only some of the more obvious questions that need to be answered about product knowledge, a major understudied area of consumer research.

The use of chronometric analysis to explore the structure of product knowledge has recently been demonstrated by Johnson and Russo (1977). The duration of a cued retrieval task identified the distance in the memory structure between the cue and the target item. This method was used to show that memory structure was largely determined by the input organization of the product information. Interestingly, some reorganization of the structure did occur, and it appeared to favor storage by attribute.

Just as important as the structure of product knowledge is the content of that knowledge. In a study of multiple decision making on a single information set, Green, Mitchell and Staelin (1977) found that subjects do not examine all the available information and more importantly, they are able to retrieve only a small portion of the examined information in memory. Examining what information is stored, in what form and how it enters memory are natural problems for response time techniques.

The importance of the content issue can be demonstrated by noting that either verbal report techniques or self administered scales have traditionally been used in most applied and theoretical research for identifying the contents of semantic memory. Of interest is how much of this reported information is based on stored information and how much on inference. While the retrieval of stored information is apt to be similar in the laboratory and in a real purchase situation, the inference processes may be very different in these two contexts.

A third use of reaction times is for examining the development of cognitive structures. Familiarity with the product class and experience with a brand have been shown to be important determinants of both purchase behavior and evaluation processes. For example, Dover and Olson (1977) have presented evidence that the predictability of the Fishbein model changes as experience with a brand changes. Psychological theory suggests that learning about a particular concept both strengthens and

adds components to a cognitive structure until it becomes unitized, i.e., may be considered a single chunk of information (Hayes-Roth, 1977). When this occurs, retrieval of information from the cognitive structure will be much quicker than at earlier stages.

In a pretest for one of our experiments (Edell and Mitchell, 1977) one group of subjects was exposed to product information. A second group received this information and, in addition, was allowed to use sample products. The resulting cognitive structure of the second group enabled the retrieval of the identical information to be completed, on the average, a full second faster than for the first group.

The previous illustrations have suggested the use of response times to measure the structure, content, or development of cognitive structures for storing product knowledge. Many of the questions raised can be answered with the simplest use of total RT's, without partitioning the process into stages or subprocesses. Others, such as determining the subprocesses involved in information acquisition, will require the use of the subtractive or additive factors methods. Although there is considerable opportunity for fruitful application of these various RT techniques to consumer research, we conclude this section with a few caveats to overenthusiasm.

Many areas of consumer behavior do not yet have process-oriented theories with the strength to support chronometric experimentation. If one cannot start without a reasonable idea of the processes being examined, no methodology, chronometric analysis included, is capable of single-handedly generating truth. Second, using response times to identify the information retrieved from memory versus that based on inferences may prove very difficult. This is especially true if retrieval times exhibit large variances both between and within individuals, as they usually do. Third, unlike artificial laboratory environments, many questions appropriate to consumer situations do not have correct answers. This will cause difficulties in adjusting RT's for the effects of a speed-accuracy trade-off. Finally, note that chronometric analysis is a sophisticated methodology developed by experimental psychologists over a period of many years. Strong behavioral and experimental expertise is required of researchers who want to benefit fully from the application of this technology to consumer behavior.

An Examination of Low Involvement Effects

Low involvement learning of advertising information was first hypothesized by Krugman (1965) over a decade ago. Since that time the concept has generated considerable speculation (e.g. Ray, 1973), but almost no empirical research. Perhaps one reason for the dearth of empirical research is that we do not have a precise definition of low involvement learning in terms of cognitive processes. One of the goals of this study is to provide such a definition.

High Involvement Learning

We might begin the task of defining low involvement by first considering what occurs in high involvement learning. Here we hypothesize that interest in the product category is high and the consumer is actively processing the information in the advertisement to reach an overall evaluation of the advertised brand. For example, individuals intending to purchase stereo components may actively seek and read advertisements to evaluate alternative components or vendors. In this situation the consumer has a goal, evaluation of alternative components, which determines the processing strategy that is used during exposure to an advertisement. Under these condi-

tions, we would expect the consumer to critically evaluate the information about the brand from the advertisement by generating thoughts or cognitive responses (Wright, 1973, Greenwald, 1968).

Low Involvement Learning

Alternatively, no evaluative processing may take place during exposure to an advertisement. This may occur, for instance, when an individual becomes involved in the photograph of a print advertisement for a brand in a product category that he or she does not intend purchasing. Alternatively, this may also occur when an external stimulus diverts attention from the processing of information contained in the advertisement. Since in both of these situations little evaluative processing is taking place, we would not consider them examples of high involvement learning.

Based on these examples, let us distinguish between two causes of low involvement processing of advertising messages: attention limitations and strategy limitations. Processing is attention limited when the advertisement does not receive enough attention for it to be fully perceived or evaluated. Following Kahneman (1973) or Norman and Bobrow (1976), attention is considered to be a limited cognitive resource. In other words, the system for processing information at any given point in time has limited capacity. Consequently, insufficient attention can be caused by a deliberate diversion of attention to other processes. For instance, an ongoing conversation with another person may divert whatever attention the individual was devoting to a radio advertisement. This type of low involvement learning has been termed distraction and has been studied by psychologists (Festinger and Maccoby, 1964) and within a marketing context (Bither, 1972). Theoretically, the diversion of attention to other processes limits the attention devoted to the processing of information from the advertisement and, consequently, reduces the number of thoughts or cognitive responses that might be generated. The usual research result under this paradigm is that more positive attitudes are formed with distraction than without. Notice that the attention limitation may be caused by either the individual or the environment.

Low involvement learning can also occur in situations of full attention, if that attention is not devoted to an evaluative processing of the advertising message. Recall the earlier example of attention devoted to the photograph in a print advertisement rather than to brand related information. We call this strategy-limited low involvement because the person processes the advertisement with other than an evaluative strategy. Under these conditions few, if any, cognitive responses about brand information from the advertisement would be generated during exposure to the advertisement with little or no resulting effect on the contents of semantic memory for the brand. However, since the individual was devoting full attention to the advertisement, a trace of the advertisement would be stored in episodic memory (Tulving, 1972). At some later point, information from the advertisement might be retrieved from episodic memory and transferred to semantic memory. However, during this transfer process, we hypothesize that the information would not be as critically evaluated as it would be if it had been received directly from the environment under conditions of high involvement. Consequently, we would expect differences in the resulting cognitive structures after exposure to the same advertisement under the two conditions. More specifically, we would expect more positive attitudes to be formed under conditions of strategy limited processing as opposed to high involvement processing.

This phenomena of strategy limited low involvement may be especially prevalent in the broadcast media where

effort is required to avoid exposure to an advertisement. In situations where the viewer has little interest in the brand being advertised, he or she may resort to nonevaluative processing instead of exposure avoidance.

Several further points should be noted. First, both attention-limited and strategy-limited low involvement can occur simultaneously. The same advertisement can experience low involvement for both reasons. Second, in neither low involvement situation is it claimed that no evaluative processing occurs. Instead, partial evaluative processing can, and probably does, occur under both attention-limited or strategy limited conditions. This partial processing, combined with repeated exposure, can lead to the learning of the advertising message. Third, the low and high involvement conditions described previously should be viewed as end points on a continuum. One example of a point along this continuum might be the case where an individual processes information from an advertisement to learn about an alternative, but he or she does not choose to make an overall evaluation of the brand at this point. Fourth, the analysis of low involvement learning presented above is constructed within an information processing framework. The emphasis is on the cognitive processes that are active during the perception of an advertisement. The subject state or even the intentions of the individual are relevant only in as much as they affect those processes.

Since attention-limited low involvement learning has been studied previously, we direct our efforts toward strategy limited low involvement learning under the condition of full attention.

Research Purpose

The purpose of this study is threefold. First, we wanted to demonstrate the nature of processing under strategy limited low involvement conditions and contrast it to processing under high involvement conditions. Second, we wanted to examine the differential effects on the resulting cognitive structures of individuals exposed to the same advertisements under these two conditions. Finally, we wanted to demonstrate the usefulness of chronometric analysis in examining the processing that takes place during exposure to advertisements.

Consequently, we wanted to design an experiment where the manipulation would be the creation of the two different involvement conditions during exposure to a number of advertisements. The difference in processes under the two conditions would be demonstrated by chronometric analysis. Subjects would be asked to verify statements which would require different mental processes and, therefore, different response times, under the two conditions. Finally, measures would be taken of the resulting cognitive structures of the brands featured in the advertisements.

The experiment should be viewed as an exploratory study whose purpose was to provide insights into the effects of low involvement processing of advertising information and the general cognitive processes involved in the learning of information from advertising.

Method

Rationale

High involvement was produced in the experiment by instructing subjects to examine the advertisements as though they were planning a purchase of the product class of the brand in the advertisement. To enforce this evaluation task, subjects were required to think aloud while making their evaluations and if they did not reach an

overall evaluation of the brand, they were prompted to do so.

The low involvement task involved evaluating the advertisements as to their ability to attract and hold attention. Since subjects would generally be unfamiliar with this task, a number of criteria were given to them which they were to use in making their evaluations. These criteria were the amount of otomotopia, assonance, alliteration, rhyme, hyperbole in the copy and the number of times the words "you" and "your" appeared. Subjects were also required to think aloud while making their evaluations and again, if they did not reach an overall evaluation of the advertisement, they were prompted to do so.

Statements were developed whose verification should require different cognitive mental processes under each experimental manipulation. Since little is known about these processes our model was necessarily simple. The high involvement group was expected to form both brand and attribute evaluations during exposure to the advertisements, therefore, the only cognitive processes required in verifying these statements should be the direct retrieval of information from semantic memory. For the low involvement group, verification of similar statements should require both retrieval of the information contained in the advertisement from episodic memory and then additional processing of this information to form attribute and brand evaluations. Therefore, the response times to these statements should be longer for the low involvement group than the high involvement group.

Similarly, the low involvement group should form evaluations with respect to the different "attention" criteria and an overall evaluation of the advertisements' ability to attract and hold attention. Verification of these types of statements should require only the direct retrieval of this information from semantic memory. The high involvement group, however, will need to retrieve the trace of the advertisement from episodic memory and then evaluate the advertisement according to the stated criteria in order to verify these statements. Therefore, the response times for the high involvement group should be longer than those for the low involvement group on these statements. An example of each type of statement is given in Table 1.

TABLE 1

EXAMPLES OF STATEMENTS

Brand Evaluation

New Car received an overall rating in the top 50% for all cars.

Attention Evaluation

The ad for New Car scores well on attention getting ability.

Attribute Evaluation

New Car is rated in the top 50% of all cars for ease of repair.
New Car is rate in the most expensive 50% of all cars.

Attention Criteria

The ad for New Car contains few rhymes.
The ad for New Car uses words like "you" and "your".

Design

An after only design was used with the main manipulation being the two different levels of involvement. Two different dependent variables were examined. The first was the response time to the four different types of statements: brand evaluations, attribute evaluations, attention evaluations and attention criteria evaluations. Advertisements for four different brands from four different product categories were used. For each brand, one brand evaluation, one attention evaluation, two attribute evaluation and two attention criteria evaluation statements were used.

The second dependent variable was the attitude toward each brand. The mean score of three bipolar evaluative scales (good-bad, dislike very much-like very much, poor quality-high quality) was used for this measure. These evaluations were obtained from a questionnaire administered after the response time statements.

Stimuli

The product categories selected for the advertisements were automobiles, pens, tires and shoes since these types of products would be familiar to our subjects. The copy for each advertisement was designed to provide factual information about the brand so that an overall evaluation of the brand along two different attributes, such as "riding comfort" for the automobile, could be made by subjects. The copy was then given to an artist who made an ink sketch of a picture to go with the advertisement. Each advertisement contained a picture, headline below the picture and then copy below the headline. The advertisements were not of professional quality. It was obvious to the subjects that they were constructed by the experimenters.

Subjects

The subjects were 30 individuals of both sexes recruited from the campus community. The sample included students, secretaries and staff members. The payment of subjects was based on the number of correct responses to a subset of response time statements. This subset includes attribute evaluations and attention criteria evaluations. Subjects received 10¢ for each correct response and lost 10¢ for each incorrect response. Any subject who earned less than \$1.00, based on the number of correct and incorrect responses, was given a dollar.

Apparatus

Response times were measured on a tachistoscope. The questions were placed at one end of the apparatus. Subjects looked into the tachistoscope at the other end, but could not see the questions until they pressed the starter pedal. This simultaneously exposed the question and started the timer. The subject answered the question by pressing either a "true" or "false" button which stopped the timer.

Procedure

Subjects were alternately assigned to low and high involvement conditions to achieve randomization and were run one at a time. At the beginning of the experiment subjects were given instructions depending on their assigned experimental condition and then given two advertisements to evaluate so they would become familiar with their respective tasks. Next they were shown the four advertisements one at a time and asked to think aloud as they did their evaluations. These verbal responses were recorded on a tape recorder.

Subjects were then directed to the tachistoscope and given instructions as to its use. They were then given ten sample statements, completely removed from the context of the experiment, to allow them to become familiar

with the operation of the machine. Prior to the experimental questions, subjects were informed that they may receive some questions that they might not be prepared to answer but that they could do so if they thought back to the advertisements. They were also told that factual statements in the advertisement were true, that they should attempt to answer all questions with equal accuracy, and were informed of the payment schedule. Finally, subjects in the high involvement task were given the criteria for evaluating each advertisement on its ability to attract and hold attention.

The statements were asked in the same order for all subjects. The order was: brand evaluation statements, attention evaluation statements, attribute evaluation statements, and attention criteria statements. Within each statement type, the order of the products was also the same: car, pen, tire and shoe.

After answering the questions on the tachistoscope, the subjects were given a questionnaire to complete and after this they were debriefed and paid. The entire procedure took approximately an hour for each subject.

Data Analysis

Response time data are almost always skewed upward (Pachella, 1974). In this experiment, the average response time over all subjects and conditions was 4.45 seconds, yet some of the observations were as large as 23 seconds. In these instances, processes other than direct retrieval or inference were being executed.

In analyzing response times with these characteristics, two approaches are commonly employed (Pachella, 1974). Either medians are substituted for means or all observations above some cutoff value are dropped. This initial analysis of the data took the second approach. All observations that were twice as long as the average were eliminated. The resulting means were also compared to the median of all the observations for a particular cell.

When two questions were used for a particular question type, these two observations were averaged to produce a single observation for a subject. If one of the response times was greater than the cutoff value, then the response time for the other question was used for that observation.

Results

Response Times

The main prediction of the experiment is the faster verification of brand evaluation statements by the high involvement group. Similarly, the low involvement group should exhibit faster verification of attention evaluation statements. The mean response times for these two types of statements confirm both predictions. These data are plotted in Figure 3. A one tailed test of the means of the brand evaluation statements (3.54 vs. 4.61) was significant ($p < 0.001$) as was the same test on the attention evaluations (4.32 vs. 3.80; $p < 0.05$).

The mean response times for attribute evaluation and attention criteria evaluations are plotted in Figure 4. As predicted, the high involvement group verified attribute evaluation statements faster (3.53 vs. 4.00) and the low involvement group verified attention criteria statements faster (5.46 vs. 4.61). A one tailed test on the means of these two statement types was significant for both the attribute evaluation statements ($p < 0.05$) and attention criteria statements ($p < 0.005$). Notice also that, in general, it took longer for both groups to verify attention criteria statements than attribute evaluation statements.

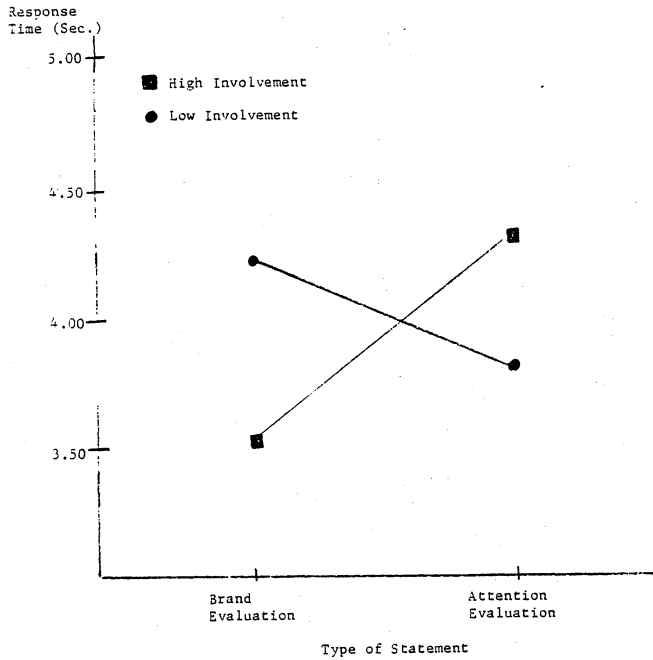


FIGURE 3. PLOT OF RESPONSE TIMES FOR BRAND AND ATTENTION EVALUATION STATEMENTS BY INVOLVEMENT LEVELS

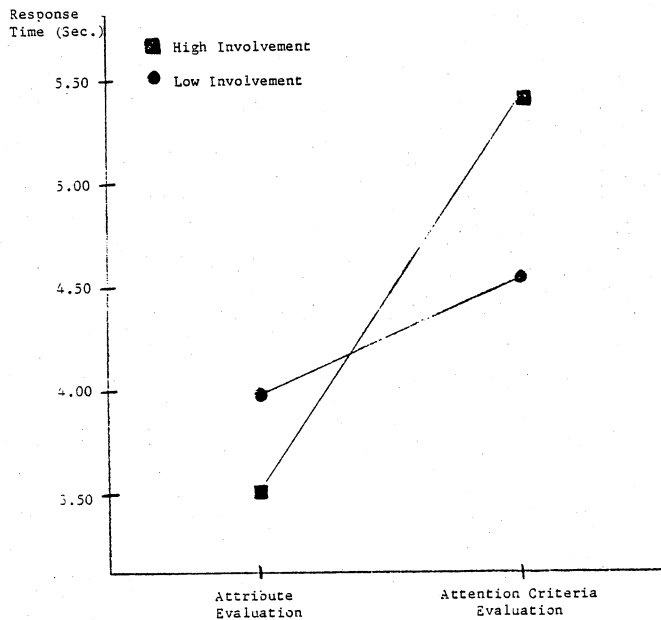


FIGURE 4. PLOT OF ATTRIBUTE AND ATTENTION CRITERIA EVALUATION STATEMENTS BY INVOLVEMENT LEVELS

Medians

The median scores for each group by statement type is given in Table 2. All of the predicted effects, except for attention evaluations, are confirmed. An examination of the data indicated that two subjects in the low involvement group were largely responsible for this result. Their response times to attention evaluation statements were consistently above 8 seconds which caused the median of this group to be almost equal to the median response of the high involvement group on this type of statement.

TABLE 2
MEDIAN BY GROUP AND STATEMENT TYPE

Group	Statement Type			
	Brand Evaluation	Attention Evaluation	Attribute Evaluation	Attention Criteria
High Involvement	3.32	4.06	3.37	5.57
Low Involvement	4.38	4.16	3.74	4.77

In general, the mean response times, after eliminating observations greater than 9 seconds, are larger than the medians. This indicates that in these cases the data is still skewed.

Correct Responses

As mentioned previously, the attribute evaluation and the attention criteria evaluation statements were structured so there was a correct response. We hypothesize that the high involvement group should be more accurate on attribute evaluation statements and on attention criteria evaluation statements, the low involvement group should be more accurate. This prediction held for the attribute evaluation statements, where the high involvement group verified a 77% of the statements correctly, while the low involvement group verified 63% correctly. This difference is significant for a one tailed test ($p < 0.05$). The high involvement group performed slightly better than the low involvement group on the attention criteria statements (55% vs. 52%), however, this difference is not significant ($p < .25$).

Attitude Toward the Brand

A 2 x 4 ANOVA on each subject's overall attitude toward the brand (Table 3) resulted in both a main effect due to group ($p < 0.01$) and product ($p < 0.001$). These main

TABLE 3
ANOVA: ATTITUDE TOWARD THE BRAND

Source	MS	d.f.	F
Between Subjects			
Group Effect	10.016	1	10.487 ^b
Subjects with Groups	0.955	28	
Within Subjects			
Product Effect	4.071	3	7.306 ^a
Product Group Interaction	0.839	3	1.5059
Product Subjects within Group Interaction	0.557	84	

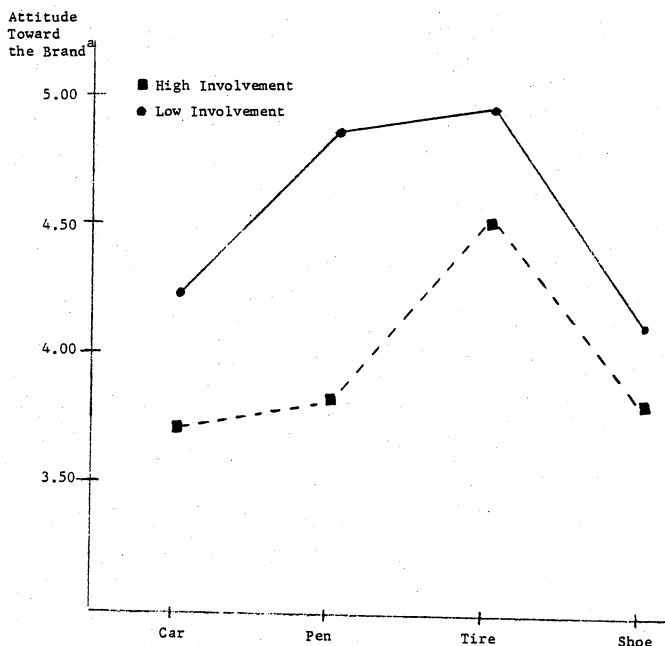
^a $p < 0.001$

^b $p < 0.01$

effects indicate a systematic difference in attitudes between groups and products. The group-brand interaction was not significant. A plot of the mean attitude scores for each brand and group is presented in Figure 5. It should be noted that the attitude toward each brand is more positive in all cases for the low involvement group.

SUMMARY AND DISCUSSION

In this paper we have discussed the use of chronometric analysis to examine internal information processing. Different approaches for using the technique were presented, along with methodological problems and finally suggestions of areas in consumer behavior where the technique might be applied.



^aLarger values indicate a more positive attitude

FIGURE 5. PLOT OF ATTITUDES TOWARD THE BRAND BY GROUP AND PRODUCT

An experiment examining the effects of high and low involvement learning of advertising information which utilized response time measures was described. Two different types of low involvement learning were hypothesized. The first, attention limited, occurs when only partial attention is directed at advertisements and has been examined previously as distraction effects. The second, strategy limited, occurs when individuals devote full attention to an advertisement, but do not process the information contained in the advertisement to reach overall evaluations.

In this study, strategy limited low involvement learning under the condition of full attention was examined. Response times were used to examine the retrieval of different types of information from memory under the two conditions. The pattern of the response times between groups for different statement types indicated that the high and low involvement groups processed information from the advertisements differently. Furthermore, the data indicated that more positive attitudes toward the brands were formed under low involvement conditions.

Problems

The preliminary results from the experiment illustrate many of the problems in using chronometric analysis in general and to examine consumer behavior specifically. First, the distribution of response times for each statement type within groups was highly skewed. Generally, observations that are considerably greater than average indicate that the subject, in that particular instance, was executing a very different process. These occurrences must be taken into account in the data analysis. In our initial analysis, response times greater than twice the overall average response time were eliminated. Second, there was some indication that the two groups may have been operating on different parts of the speed-accuracy operating characteristic. The low involvement group had slightly lower accuracy on the attention criteria statement than the high involvement group although it had been hypothesized that they should have had greater accuracy. This may be an indication that the response times for this group had a downward bias. Finally, the results

indicated that a model of the process that is more sophisticated than the one initially postulated may be required. For instance, more time was required to retrieve attention criteria information than attention criteria information. Our simple model did not provide an explanation of these differences.

In summary, the experiment indicated that low involvement learning of advertising can occur and that in our experiment more positive attitudes were formed with low involvement learning than with high involvement learning. We are now in the process of doing additional data analysis to understand why more positive attitudes were formed under low involvement conditions and also examining the amount of learning of information under each condition.

References

- J. R. Anderson and H. Bower, Human Associative Memory, (New York: John Wiley and Sons, 1973).
- J. R. Bettman, An Information Processing Theory of Consumer Choice, (Reading, Mass.: Addison-Wesley, 1978), forthcoming.
- S. W. Bither, "Effects of Distraction and Commitment on the Persuasiveness of Television Advertising," Journal of Marketing Research, 9 (1972), 1-5.
- W. G. Chase, "Elementary Information Processes," in William K. Estes (ed.), Handbook of Learning and Cognitive Processes Vol. 5, (Hillsdale, N.J.: Lawrence Erlbaum, 1978), in press.
- W. G. Chase and H. H. Clark, "Mental Operations in the Comparison of Sentences and Pictures," in L. Gregg (ed.), Cognition in Learning and Memory, (New York: John Wiley and Sons, 1972).
- R. W. Chestnut and J. Jacoby, "Consumer Information Processing: Emerging Theory and Findings," in Arch G. Woodside, et. al. (eds.), Consumer and Industrial Buying Behavior, (New York: Elsevier North-Holland, Inc., 1977).
- H. H. Clark and W. G. Chase, "On the Process of Comparing Sentences Against Pictures," Cognitive Psychology, 3 (1972), 472-517.
- A. M. Collins and M. R. Quillian, "Retrieval Time From Semantic Memory," Journal of Verbal Learning and Verbal Behavior, 8 (1969), 240-248.
- F. C. Donders, "On the Speed of Mental Processes," Acta Psychologica, 30 (1969), 412-431. Translated from the original by W. G. Koster from Onderzoekingen gedaan in het Psychologisch Laboratorium der Utrechtsche Hoogeschool, 1868, Tweede reeks, II, 92-120.
- B. A. Doshier, "Retrieval of Sentences from Memory: A Speed-Accuracy Study," Cognitive Psychology, 8 (1976), 291-310.
- P. A. Dover and C. Olson, "Dynamic Changes in the Expectancy-Value Attitude Model as a Function of Multiple Exposures to Product Information," in B. A. Greenberg and D. A. Ballenger (eds.), Contemporary Marketing Thought, (Chicago: American Marketing Association, 1977), 455-460.
- J. Edell and A. Mitchell, "The Mediating Effect of Product Class Familiarity Upon Cognitive Responses to Advertising," Working Paper, Graduate School of Industrial Administration, Carnegie-Mellon University, 1977.

- L. Festinger and N. Maccoby, "On Resistance to Persuasive Communications," Journal of Abnormal and Social Psychology, 68 (1964), 359-366.
- R. Green, A. Mitchell and R. Staelin, "Longitudinal Decision Studies Using a Process Approach: Some Results From a Preliminary Experiment," in B. A. Greenberg and D. A. Ballenger (eds.), Contemporary Marketing Thought, (Chicago: American Marketing Association, 1977), 461-466.
- A. G. Greenwald, "Cognitive Learning, Cognitive Response to Persuasion, and Attitude Change," in A. G. Greenwald, T. C. Grock and T. M. Ostrom (eds.), Psychological Foundations of Attitudes, (New York: Academic Press, 1968).
- B. Hayes-Roth, "Evolution of Cognitive Structures and Processes," Psychological Review, 84 (1977), 260-278.
- E. J. Johnson and J. E. Russo, "The Organization of Product Information in Memory Identified by Recall Times," by H. Keith Hunt (ed.), Advances in Consumer Research, Vol. 5, Association for Consumer Research, forthcoming.
- D. Klahr and J. G. Wallace, Cognitive Development, (New York: John Wiley and Sons, 1976).
- H. E. Krugman, "The Impact of Television Advertising: Learning Without Involvement," Public Opinion Quarterly, 29 (1965), 344-356.
- R. E. Nesbitt and T. Decamp Wilson, "Telling More Than We Can Know: Verbal Reports on Mental Processes," Psychological Review, 84 (1977), 231-259.
- D. A. Norman and D. E. Rumelhart, Explorations in Cognition, (San Francisco: W. H. Freeman, 1975).
- J. C. Olson, "Theories of Information Encoding and Storage: Implications for Consumer Research," in Andrew Mitchell (ed.), The Effect of Information on Consumer and Market Behavior, (Chicago: American Marketing Association), forthcoming.
- R. G. Pachella, "The Interpretation of Reaction Time in Information-Processing Research," in Barry H. Kantowitz (ed.), Human Information Processing: Tutorials in Performance and Cognition, (Hillsdale, N.J.: Lawrence Erlbaum Associates, 1974).
- R. W. Pew, "The Speed-Accuracy Operating Characteristics," Attention and Performance II, Acta Psychologica, 30 (1969), 16-26.
- M. R. Quillian, "The Teachable Language Comprehender: A Simulated Program and Theory of Language," Communications of ACM, 12 (1969), 459-476.
- M. Ray, "Marketing Communication and the Hierarchy of Effects," Working Paper, Marketing Science Institute, Cambridge, Mass., November 1973.
- S. Sternberg, "The Discovery of Processing Stages: Extensions of Donder's Method," Attention and Memory Performance II, Acta Psychologica, 30 (1969), 421-457.
- S. Sternberg, "Memory Scanning: Mental Processes Revealed by Reaction Time Experiments," American Scientist, 57 (1969), 421-457.
- E. Tulving, "Episodic and Semantic Memory," in E. Tulving and W. Donaldson (eds.), Organization of Memory, (New York: Academic Press, 1972), 385-387.
- E. Tulving and G. H. Bower, "The Logic of Memory Representations," in Gordon H. Bower (ed.), The Psychology of Learning and Motivation, 8, (New York: Academic Press, 1974).
- P. Wright, "The Cognitive Processes Mediating Acceptance of Advertising," Journal of Marketing Research, 10 (1973), 53-62.

W. Fred van Raaij, Tilburg University
Kassaye Wandwossen, Tilburg University

Abstract

Motivation-need theories are reviewed, their implications to consumer behavior investigated, and the various findings and concepts integrated in formulating a model of generic choice prediction.

product attributes. Absence of necessary attributes gives rise to dissatisfaction, while the presence of motivating attributes leads to satisfaction.

Instincts and Needs

In the mainstream of Darwinian evolutionary theory, certain behavioral tendencies are innately built into organisms for survival of the individual and thus the species. William James (1890) and William McDougall (1923) made lists of instincts that were seen as main-springs of all kinds of behaviors, simple and complex. Later, Murray (1937) made another classification of human needs. Murray, however, distinguished a directional aspect and an arousal component that actually kicks the behavior off and that can be motivated in a number of ways. Needs, in Murray's concepts, are hypothetical constructs directing behavior toward certain goals, or end states. Classifications of needs, as provided by McDougall or Murray look similar to classifications of elements in chemistry, but lack their strictly defined structure and usefulness. A structural principle is needed to explain the dynamic interactions of needs and their fulfillments.

Maslow's Need Hierarchy

Abraham Maslow (1965) postulates that needs are hierarchically structured and that needs low in the hierarchy must be fulfilled before need higher in the hierarchy become salient. Interrelationships between needs are specified, which are missing in McDougall's and Murray's systems. According to Maslow, the physiological needs (e.g., hunger, thirst) come first, followed by security needs, social needs (affiliation), self-esteem needs (recognition), and finally self-actualization needs. Recently, the need to know and to understand, and aesthetic needs are added to the list (Maslow, 1970). The need to know and to understand is comparable to Berlyne's (1963) epistemic behavior.

Maslow's basic needs (Maslow, 1943, 1965, 1970) are thought to be structured in such a way that the satisfaction or gratification of the lower-order needs leads to the activation of the next higher-order need in the hierarchy. This is the gratification/activation principle. The other principle is the deprivation/domination principle, which states that the most deficient need is the most important need. A distinction is sometimes made between deficiency and growth needs. Needs for belongingness, love, and self-actualization are referred to as growth needs; the others are deficiency needs. To achieve growth needs, deficiency needs must first be satisfied. This may be compared with Herzberg's (1966) hygienic and motivating factors in his job satisfaction theory. A deficiency in the hygienic factors creates dissatisfaction, while fulfillment of these factors does not create satisfaction. The motivating factors, when fulfilled, give rise to job satisfaction. Job satisfaction, and probably also consumer satisfaction, is not measurable on a simple bipolar scale but consists of two more or less independent (sets of) factors. In consumer research, we may distinguish between necessary product attributes (hygienic factors) and motivating

Evaluation of Maslow's Need Hierarchy

Despite its vagueness and lack of adequate empirical support (Wahba and Bridwell, 1976), Maslow's need hierarchy has influenced the work of numerous psychologists (Argyris, 1964; Clark, 1960; Dichter, 1964; Leavitt, 1964; McGregor, 1960; Schein, 1965). Nonetheless, the findings remain largely controversial; and an evaluation of interdisciplinary approaches is rare (Jacoby, 1976). A number of factors seem to have favored the appeal of Maslow's need hierarchy, while the lack of foresight among researchers and the absence of standardized measurement techniques seem to have forestalled the comprehensive evaluation of the interdisciplinary approaches. 1. Maslow's approach is a theory of motivation, in that it links basic needs-motives to general behavior (Wahba and Bridwell, 1976). In other words, the basic needs/motives are linked to behavior through a theory of motivation which asserts that (i) deprivation is followed by gratification; (ii) less potent needs emerge upon the gratification of the more preponderant ones (Maslow, 1970); (iii) and it is a dynamic process where deprivation is hypothesized to lead to domination, which leads to gratification that culminates in the activation of the next higher order need in the echelon. 2. At times, the findings have been used to support two apparently contradictory hypotheses. For instance, Maslow (1965) postulated that (i) gratification of the self-actualization need results in an increase of its importance rather than a decrease, and also that (ii) a long-time deprivation of a need may create a fixation for that need. Maslow noted the exception to his model; that, it is possible for higher-order needs to emerge not after gratification of the next-lower need, but after long-time deprivation (Maslow, 1970). The state of affairs remains that Maslow's need hierarchy, and his propositions regarding gratification and activation, especially in the self-actualization stage, remain controversial. His need hierarchy is by no means definitive, and is rather out of focus in comparison with the role of learning, perception, values, and expectations in human behavior (Atkinson, 1964).

Alderfer (1972) points out that satisfaction with regard to some environmental and job characteristics are studied rather than satisfaction with the postulated needs. Maslow initially postulated that high satisfaction or dissatisfaction is given high ranked importance (Maslow, 1965). Contrary to what is postulated by Maslow, high job satisfaction rather than deprivation is correlated with importance (Dachler and Hulin, 1969). In another study, again, contrary to what Maslow hypothesized, Mobley and Locke (1970) concluded that extreme satisfaction and dissatisfaction depend on the importance attached to them, and not importance determining satisfaction and dissatisfaction.

Some support has been found for Maslow's (1965) deficiency and growth needs in studies that compared executives and workers in an organization. In these studies higher-order (growth, actualizing) needs are judged to be more important for top executives than for underprivileged workers (Davis, 1946; Pellegrin and Coates, 1957). In a deprived environment, lower-order needs

(existence hygienic) needs seem to be more important than higher order need, (Cofer and Appley, 1964; Porter, 1961, 1962; Porter and Mitchell, 1967). By implication, the rating of importance of job satisfaction seems to be positively related to the level of the job one holds (Porter, 1961; Porter and Mitchell, 1967) or "that the deprivation domination principle may only be operative in the case of the deprivation of the lower-order needs, especially physiological needs" (Wahba & Bridwell, 1976, p. 231).

Notwithstanding the above conclusions, the concept of deprivation/domination seems to have little or no effect on the behavior of consumers in relatively affluent societies for a number of reasons. (1) The daily purchases are mostly over and above what is (basically) needed. (2) Until the time that the law of diminishing returns sets in or depleting raw material resources make "abundant" consumption difficult, there is a "need" to buy and possess more. Instead, consumption is influenced by relative deprivation compared with "relevant other consumers". This relative deprivation may trigger the dominance of the desire "to keep up" with the reference group.

Satisfaction/Dissatisfaction

Consumer satisfaction/dissatisfaction studies will benefit from the two-factor need theories (Maslow, 1965; Alderfer, 1969; Herzberg, 1966). Brands possess two types of attributes. The first type of attributes (inhibitors) give rise to dissatisfaction, if their level is below a certain threshold. A car that is insufficiently safe causes dissatisfaction, while no satisfaction is derived from a car that is sufficiently safe. The second type of attributes (facilitators) give rise to satisfaction, if their level is above a certain threshold. Similar to the deprivation/domination principle, the presence of inhibitors causes dissatisfaction and (extending the above principle) this dissatisfaction cannot be compensated for by facilitators. If no inhibitors are present, a "zero point" has been reached. Consumer satisfaction can only be obtained through the absence of inhibitors and the presence of facilitators.

The distinction between inhibitors and facilitators has its analogy in consumer decision making. In terms of decision rules, the first type of attributes (inhibitors) elicit the conjunctive decision rule to eliminate brands with inhibiting (below threshold) values on certain attributes. The second type of attributes (facilitators) elicit the disjunctive decision rule to select brands with facilitating (above threshold) values on other attributes. The conjunctive rule must occur before the disjunctive rule.

Jacoby (1976) emphasizes the applicability of Herzberg's (1966) two-factor model for the study of consumer satisfaction, which may be compared to a simple choice heuristic: the sequence of conjunctive and disjunctive information processing (Van Raaij, 1977, p. 23-26). Some problems exist, however, in applying Herzberg's (1966) two-factor model in consumer satisfaction research:

1. In the decision process the consumer will avoid brands that give rise to dissatisfaction through the application of the conjunctive decision rule. Dissatisfaction may only occur after an incorrect application of the conjunctive rule, or after using incomplete or deceptive information.

2. As Jacoby (1976) points out, Herzberg's propositions as well as the findings cited before are involved with the determinants of satisfaction/dissatisfaction and not with performance. Therefore, they cannot be directly extended to a purchase situation that involves a combination of dichotomies involving purchase behavior-satisfaction and purchase behavior-dissatisfaction. As a solution, he proposes another behavior-satisfaction dimension orthogonal to the facilitator-inhibitor di-

mension (Jacoby, 1971).

Motivation Models

Cognitive motivation models fall into three broad categories: equity, need achievement and expectancy-value models (see Table 1). The commonality of the models is that the units of framework we present are of cognitive, subjective nature, and that they include hypothetical constructs as perceived equity, need achievement, expectation, and values. A similar, but shorter, review of the three categories is given in Van Raaij (1976).

Equity

The concept of equity may be explicitly stated as the even exchange of values such that what is received is presumed to be equal to what is given (Adams, 1965). Equity operates within a range, with a lower and upper limit. Inputs ("what is given") are defined as "what a person perceives as his contributions to the exchange for which he expects a just return" (Walster and Walster, 1975). Apparently, equity theory may be applied to social relations such as amnagement-worker and seller-buyer. Further, the concept of power seems to be related to perceived and subjective equity.

It is our hypothesis that the ranges of equity (upper and lower limits) may well be measured by the expectancy-value type of model (Table 1) for two reasons: (1) The expectancy component of the model is general, comprehensive and brand specific. (2) The expectancy component handles expectations about equity as compared with "relevant others". We return later to the discussion of how consumer behavior is motivated by perceived inequity or a disparity between the desired and actual state. The application of the equity concept of consumer behavior may be restricted to some aspects of consumption.

TABLE 1
Motivation Models

Type	Formulae	Explanation
1. Equity-expectancy		
a. job performance: (Vroom, 1964)	$MF = E \times V$	MF = Motivational force E = Expectancy force of achieving desired outcome V = Value of desired outcome
b. attitude:	$A_o = PI \times VI$	A = Attitude PI = Perceived Instrumentality VI = Value importance
2. Need-achievement:	$T_s = M_s \times P_s \times I_s$ $T_f = M_f \times P_f \times I_f$ $I_f + P_s = 0$ and $P_s + P_f = 1$ and $P_s + I_s = 1$ $T_a = T_s - T_f$ $T_a = (M_s - M_f) \{P_s - P_s^2\}$	T_s = Strength of motivation to achieve success T_f = Strength of motivation to avoid failure M_s = Motive or need to achieve success M_f = Motive or need to avoid failure I_f = Incentive value of failure I_s = Incentive value of success P_f = Probability of failure P_s = Probability of success

Such equity-based motivational forces include sensitivity of consumers toward primarily price, time and effort expended (e.g., Gabor and Granger, 1966). However, it has to be pointed out that (1) it is not prior known how equity is created and what its upper and lower limits are; (2) promotional activities make the equity relation relative and situation-affected, depending on whether the purchase has been prompted by a deal or not, for instance; (3) consumers tend to "satisfice" (March and Simon, 1958), and do not necessarily maximize as implied in the equity concept (e.g., Pritchard, 1969).

To summarize, equity is useful in two ways for consumer research. First, the inequity of the seller-buyer relation may give rise to consumer dissatisfaction and the motivation to restore equity. Armstrong (1976) provides some examples of the restoration of equity between consumers and marketers. As with Herzberg's (1966) two-factor model, an equitable relation as perceived by the consumer prevents the elicitation of dissatisfaction. On the other hand, an inequitable relation is a necessary but not a sufficient prerequisite for consumer satisfaction. Second, the equity relations holds for the consumer with regard to "relevant others" (reference groups). Here, an inequitable relation motivates the consumer to restore equity, that means he is motivated to bring his consumption level and pattern into agreement with that of his reference group. Generally, this results in an increase in consumption expenditures, as aspirations and expectations become higher (e.g., Duessenberry's 1949) "relative income hypothesis" and Katona, Strumpel and Zahn's (1971) "rising aspirations and affluence").

Need Achievement

The concept of need achievement (McClelland, 1961) is basically another variation of the expectancy-value approach. Need achievement resembles Maslow's (1970) self-actualization motive in a number of ways. The main difference, however, is that it includes the probability of attaining a goal and a probability of failure. The need achievement model (Table 1) attributes the strength of motivation to the cognitive expectation that the action will result in the consequence. Stated somewhat differently, the tendency to engage in an activity is determined by the desired goal of the action. Therefore, the outcome or consequence has attraction or value to the individual. Assuming that $I_f = -P = 1$ (a particular outcome is either a success or a failure). As indicated in Table 1, the summation of T_s and T_f provides the tendency or motive to achieve (T_a), which may be derived from the given algebraic relationship: $T_a = (M_s - M_f) / (P_s - P_f)$.

The need achievement concept is only applicable in cases where the consumer perceives some risk of failure. Two ways exist to increase the tendency to achieve (T_a): (1) Increase the "approach" tendency (T_s) by making the product more attractive, and (2) decrease the "avoidance" tendency (T_f) by reducing perceived risk (see Roselius (1971) for possible risk relievers). Again, a congruence with Herzberg's two-factor model can be observed. The avoidance tendency (T_f) may be related to some unsatisfactory product attributes (inhibitors) and the approach tendency (T_s) may be related to other, satisfactory product attributes (facilitators). The only difference is that in the approach-avoidance paradigm unsatisfactory attributes can be compensated by satisfactory product attributes. As Schewe (1973) points out, "The greatest problem appears to be determining a valid and reliable measure of the need achievement construct" (Schewe, 1973, p. 33). In addition, achievement needs are not operating in all purchase situations. P and I may have low levels and, hence, result in low T_s levels of the strength of the motive to achieve success (T_a). He concludes that further research is necessary to find its true potential as a determinant of consumer behavior (Schewe, 1973).

Finally, contrary to the postulate of need achievement, motives are not stable behavioral dispositions, though they may well be partly learned. Also, it must be emphasized that it is not necessary for motives in general to operate after they are aroused by the presence of the incentives or situational cues that have been associated with the incentive (see Campbell and Pritchard, 1976, pp. 112-14).

Expectancy-Value Models

The basis for expectancy models has been made by Tolman (1932) and Lewin (1938). It seems to be influenced more by Lewin's field theory in that it involves the perceptual analysis of (1) alternatives with their (2) desirabilities and (3) expectancies, and their (4) outcomes in the immediate psychological field.

Many psychological theories come under the label of expectancy-value models: subjective expected utility theory (Edwards, 1954), social learning theory (Rotter, 1954), motivation theory (Atkinson, 1964), and attitude theories (e.g., Rosenberg, 1956; Fishbein, 1967). An overview of these theories can be found in Van Raaij (1977).

Expectancy theory states that the desire or motive to engage in a certain behavior is a composite of the expected outcome of that behavior and the value or evaluation of that behavior. As can be seen from Table 1 the motivational force to engage in a particular behavior, as applied in organizational psychology, is a function of the four factors stated above (Vroom, 1964; Graen, 1969; Porter and Lawler, 1968; Campbell, Dunnette, Lawler and Weick, 1970). However, it needs to be pointed out that the expectancy concept is not without questions. We will try to find answers to these questions before we apply this concept to consumer motivation.

First, the extension of the concept of "evoked set" (Howard and Sheth, 1969) seems to provide an answer to the question of how many alternatives, as well as type of alternatives are considered by the individual. Second, the desirability or attractiveness of the alternatives is a function of the probability that the alternative possesses a certain attribute times the evaluation of that attribute on a bipolar favorable-unfavorable scale. Third, Jacoby (1976) emphasizes not to overlook that the desired outcomes of a behavior are influenced by "motivational inputs". A distinction is made between input and output. "Outputs or outcomes refer to the primary functional aspects of the alternatives in the product set; they are the basic purpose for buying and using the product. Inputs, on the other hand, are those motivational forces other than perceived functional consequences which influence the selection of one specific behavioral alternative over the other available alternatives" (Jacoby, 1976, p. 1049).

If the functional goal of the purchase of a car is its service and economy, say, as opposed to status or a combination of all these, then these consequences constitute the desired outcomes. The inputs such as advertising, availability of deals, past satisfaction with the product, referred to as "antecedents" (Jacoby, 1976), may induce the consideration of one brand over another. Jacoby suggests the partitioning of the evaluation component into input (or antecedent) and output (or consequent) "values". Unfortunately, this is only at the brand level, and even at that level, it fails to resolve a number of questions raised earlier. Although Jacoby's revision make the traditional models more comprehensive and richer in their construct composition, some drawbacks have to be mentioned: (1) It fails to answer how and why an individual becomes motivated to consider certain outcomes or consequences. (2) It ignores the interdependency between product and brand, that is, the desire to consider a product class and then to engage

in brand(s) selection. (3) It fails to note that some repetitive buying behavior is influenced by simple S-R relationships, or may even be stochastic, making motivational models too elaborate or irrelevant for this kind of buying behavior.

Our interest in motivational models of consumer behavior is mainly at the product class level (generic choice) but has also implications for the brand level (specific choice).

Motivational Model for Product Choice

We think that motivational models are especially useful for the generic choice (among product classes) and less useful for the specific choice (within product classes). For the latter case, multi-attribute attitude and preference models may hold better predictions for brand choice within the product class. In most cases, the generic choice is more important and critical for the consumer; however, this seems to be a neglected area of research in marketing. We think that the generic choice, whether to buy a car or to go on a vacation, for example, has more relevance for general economic policy, consumer education, and also for marketing mix decisions. Between the generic and specific choice, a "modal choice" or method choice can be distinguished in many cases (see, for instance, Sheth (1975) for travel mode selection).

Figure 1 gives the sequence of the three choice levels as they occur in consumer decision making regarding travel. The product choice is the first to be made. Sub-

FIGURE 1
Sequence of Choices in Consumer Decision Making

- A. generic choice (travel vs. other product classes)
- B. modal choice (airline vs. train)
- C. specific choice (American vs. United)

sequently, a selection of a modal or method within the product class is made. Then, within the mode, the consumer selects a specific brand.

Sheth (1975) distinguishes five utility needs. These utility needs can be seen as the basic needs that products satisfy. In the generic choice process, consumers compare the product classes on their ability to satisfy the basic needs. We may also conceive these utility needs as the basic dimensions of motivation. In all cases, consumers want to reach certain goal states and the products are instrumental in reaching the goals.

The five motivational dimensions are (Sheth, 1975): (1) functional motives, (2) aesthetic-emotional motives, (3) social motives, (4) situational motives, and (5) curiosity motives. 1. Functional motives are related to the technical functions the product performs. The combination of product attributes forms the total functional utility of a product. 2. Aesthetic-emotional motives are the style, design, luxury, and comfort of a product (class). These motives are not only important for the specific (brand) choice but also for the generic (product) choice. The product class is evaluated in terms of the fundamental values of the consumer in the emotive areas of fear, social concern, respect for quality of life, appreciation of fine arts, religion, and other emotional feelings. Thus, it may be contended that individuals tend to select those product classes that match with their life styles and enable them to express their fundamental values. 3. Social motives are related to the impact that consumption makes on relevant others. Status, prestige, and esteem may be derived from the possession and usage of products and their con-

spicuous features. Some products are selected for their conspicuousness only ("conversation pieces"), sometimes in combination with aesthetic motives. 4. Situational motives are not motives in the sense of long-term desires to reach a certain goal. The selection of a product may be triggered by situational determinants such as availability, price discount, and/or accessibility. These situational factors apply usually for a specific brand or type. The brand choice is usually made in these cases without a careful evaluation of the product class(es). 5. Curiosity motives are motives that are supposed to prompt trials of new and/or innovative products. The consumer may try a new product; however, his repeat-purchase may be independent of such trials.

Choice Modal Prediction

It has to be emphasized that the motivational model suggested here is mainly applicable to consumers' product choices, involving large financial outlays or high perceived social and/or physical risk. Repetitive brand or product choice triggered by depletion of stock is not relevantly described and predicted by our motivational model. Further, note that within each of the five motivational dimensions subclasses exist for different product classes. For example, the generic choice between a refrigerator and a TV set involves different functional utilities-cooling food versus entertainment/information, respectively. In the generic choice process, the consumer essentially compares products on a different set of dimensions for each product, while in the specific choice process the same set of dimensions apply for all brands within the product class. The consumer necessarily, unlike in the specific choice situation, has to "compute" an overall utility for each product class to see whether it satisfies a number of motives, and then compare these overall utilities to make a final choice. To depict these relationships, we suggest a straightforward multi-attribute model. In this model (eq. 1), overall preference or total utility a product class satisfies and the evaluation of these motives may be written as,

$$U_j = \sum_{i=1}^m M_{ij} \times V_i \dots\dots (1)$$

U_j is the utility of product class j that satisfies m motives (M_{ij}), and V_i is the evaluation of the m motives on a favorable-unfavorable scale. M_{ij} can be thought of as a vector of probabilities that the product class j satisfies a specific motive i . These probabilities are strictly zero or above zero, and therefore, only positive. This composite measure, U_j , is expected to cover the five dimensions outlined above. Thus the behavior or behavioral intention (BI) of the consumer equals the maximum of U_j ($j=1, \dots, m$) if m product classes are considered.

$$B = BI = \max U_j \quad (j=1, \dots, m) \dots\dots (2)$$

Usually, the number of product classes is smaller than the number of brands in the specific choice situation.

Individual consumers differ not only in their evaluation of motivational dimensions (V_i), but also in the saliency of these dimensions over time. Recent gratification of a motivational dimension may lead to a decrease in the evaluation of that motivational dimension. This is especially true for the functional, social, and curiosity motives. Lack of gratification of a motivational dimension increases the evaluation of that motive (the deprivation/dominance principle).

Basically, motives are "means-end beliefs" (Tolman, 1932). That is, there is cognitive association between a specific product class or the buying of a certain product from a set of product classes, and the expectation that the product contributes to the attainment of a goal or the satisfaction of a motivational dimen-

sion. This expectation is thought to be a subjective probability. Parenthetically, the strength of motives or motivational dimensions is largely determined by cultural and life history factors. Further, motives become salient if a disparity exists between a desired goal state and the actual state on a motivational dimension. A desired state is triggered in the comparison of one's own position and the position of "relevant others" on the various functional as well as non-functional utility dimension(s). Equity theory (Adams, 1965) predicts that differences in the input/output ratio bring about a change in the desired goal state. Need-achievement theory (McClelland, 1961) attributes the strength of motivation to the cognitive expectation that the action will result in the consequence. The origin of motivation is external in equity theory (reference groups) and internal in need-achievement theory. We assume that an intermediate disparity between desired and actual state of the individual has the strongest effect on motivation. For a small disparity an assimilation effect is thought to occur; the disparity is rationalized away. For a large disparity, a contrast effect is more likely; the disparity is too great to be bridged by the acquisition of a product. The desired goal state is perceived as unattainable in this case. This curvilinear relationship between motive strength and disparity may be effectively compared with Berlyne's (1963) exploratory behavior theory and the level of arousal potential.

Operationalization

The elicitation of the motivational dimensions can be done in two subsequent pilot surveys constituting depth interviews and other non-attributive methods. In the depth interview method no particular forms and orders of motives should be elicited with the help of probing questions, incomplete sentences and the Kelly grid method. Such elicited motives constitute a listing of the relevant needs or motives applicable to a specific situation. In non-attributive method the researcher has to start with a listing of these possible motives and request the consumer to indicate the ones he considers salient.

Summary

Motivation-need theories are reviewed, their implications to consumer behavior investigated, and the various findings and concepts integrated in formulating a model of choice prediction.

In our opinion, motivational models are highly relevant for the generic choice process, while multi-attribute attitude models are relevant for the specific choice process. A motivational model for the generic choice is proposed, with the notion that consumer behavior in its various ramifications (i.e., from the consumer, economist, social marketer, etc., perspectives) can be better understood from the analysis of generic choices.

References

- J. S. Adams, "Inequity in Social Exchange," in L. Berkowitz (ed.), Advances in Experimental Social Psychology, 2, New York: Academic Press, 1965.
- C. P. Alderfer, "An Empirical Test of a New Theory of Human Needs," Organizational Behavior and Human Performance, 4a(May 1969), 142-75.
- C. P. Alderfer, Existence, Relatedness and Growth, New York: The Free Press, 1972.
- C. Argyris, Integrating the Individual and the Organization, New York: Wiley, 1964.
- J. S. Armstrong, "Role Conflict in Marketing or the Ox and the Fox," Oslo: Symposium, "New Directions in Marketing," 1976.
- J.W. Atkinson, An Introduction to Motivation, New York: Van Nostrand, 1964.
- L. Berkowitz, "Social Motivation," in G. Lindzey and E. Aronson (eds.), Handbook of Social Psychology, Reading, MA: Addison-Wesley, Vol. 3, 1969 (second edition).
- D. E. Berlyne, "Motivational Problems Raised by Exploratory and Epistemic Behavior," in K. Sigmund (ed.), Psychology: The Study of a Science, New York: McGraw-Hill, Vol. 5, 1963, 284-264.
- P. Blau, Exchange and Power in Social Life, New York: Wiley, 1964.
- R. E. Burnkrant, "A Motivational Model of Information Processing Intensity," Journal of Consumer Research, 3, (June 1976), 21-30.
- J. P. Campbell, M. D. Dunnette, E. E. Lawler and K. E. Weick, Managerial Behavior, Performance, and Effectiveness, New York: McGraw-Hill, 1970.
- J. P. Campbell and R. D. Pritchard, "Motivation Theory in Industrial and Organizational Psychology," in M. D. Dunnette (ed.), Handbook of Organizational Psychology, Chicago: Rand McNally, 1976.
- J. B. Clark, "Motivation in Work Groups: A Tentative View," Human Organization, 13, (1960-61), 198-208.
- C. N. Cofer and M. H. Appley, Motivation: Theory and Research, New York: Wiley, 1964.
- H. P. Dachler and C. L. Hulin, "A Reconsideration of the Relationship Between Satisfaction and Judged Importance of Environment and Job Characteristics," Organizational Behavior and Human Performance, 4, (August 1969), 252-66.
- W. A. Davis, "The Motivation of the Underprivileged Worker," in W. F. Whyte (ed.), Industry and Society, New York: McGraw-Hill, 1946.
- D. Dichter, Handbook of Consumer Motivations: The Psychology of the World of Objects, New York: McGraw-Hill, 1964.
- J. S. Duessenberry, Income, Saving and the Theory of Consumer Behavior, Cambridge, MA: Harvard University Press, 1949.
- W. Edwards, "Probability Preference in Gambling," American Journal of Psychology, 67, (1954), 441-52.
- J. F. Engel, D. T. Kollat and R. D. Blackwell, Consumer Behavior, New York: Holt, Rinehart and Winston, 1973 (second edition).
- G. Fennel, "Motivation Research Revisited," Journal of Advertising Research, 15(June 1975), 23-8.
- M. A. Fishbein, "Attitude and the Prediction of Behavior," in M. A. Fishbein (ed.), Readings in Attitude Theory and Measurement, New York: Wiley, 1967, 477-92.
- A. Gabor and C. W. J. Granger, "Price as an Indicator of Quality: Report on an Inquiry," Economica, 33(February 1966), 43-70.

- G. B. Graen, "Instrumentality Theory of Work Motivation: Some Experimental Results and Suggested Modifications," Journal of Applied Psychology Monographs, 53(April 1969) part 2.
- D. T. Hall and K. E. Nougain, "An Examination of Maslow's Need Hierarchy in an Organizational Setting," Organizational Behavior and Human Performance, 3(February 1968), 12-35.
- F. Herzberg, Work and the Nature of Man, Cleveland, OH: World Publishing Company, 1966.
- J. A. Howard and J. N. Sheth, The Theory of Buyer Behavior, New York: Wiley, 1969.
- W. James, The Principle of Psychology, New York: Holt, 1890 (2 volumes).
- J. Jacoby, "A Multi-indicant Approach for Studying New Product Adopters," Journal of Applied Psychology, 55, (August 1971), 384-8.
- J. Jacoby, "Consumer and Industrial Psychology: Prospects for Theory Corroboration and Mutual Contribution," in D. M. Dunnette (ed.), Handbook of Organizational Psychology, Chicago: Rand McNally, 1976.
- G. B. Katona, B. Strumpel, and E. Zahn, Aspirations and Affluence: Comparative Studies in the United States and Western Europe, New York: McGraw-Hill, 1971.
- A. K. Korman, The Psychology of Motivation, Englewood Cliffs, NJ: Prentice-Hall, 1974.
- E. E. Lawler, Pay and Organizational Effectiveness: A Psychological View, New York: McGraw-Hill, 1971.
- H. J. Leavitt, Managerial Psychology, Chicago: The University of Chicago Press, 1964 (second edition).
- K. Lewin, The Conceptual Representation and the Measurement of Psychological Forces, Durham, NC: Duke University Press, 1938.
- J. G. March and H. Simon, Organizations, New York: Wiley, 1958.
- A. H. Maslow, "A Theory of Human Motivation," Psychological Review, 50(1943), 370-96.
- A. H. Maslow, "Higher and Lower Order Needs," in C. L. Stacey and M. F. DeMartino (eds.), Understanding Human Motivation, Cleveland/New York: The World Publishing Company, 1965.
- A. H. Maslow, Motivation and Personality, New York: Harper & Row, 1970, (second edition).
- D. C. McClelland, The Achieving Society, Princeton, NJ: Van Nostrand, 1961.
- W. McDougall, Outline of Psychology, Boston: Scribner's, 1923.
- D. McGregor, The Human Side of the Enterprise, New York: McGraw-Hill, 1960.
- W. H. Mobley and E. A. Locke, "The Relationship of Value Importance to Satisfaction," Organizational Behavior and Human Performance, 5(September 1970), 463-83.
- H. A. Murray, "Facts Which Support the Concept of Need or Drive," Journal of Psychology, 3(1937), 27-42.
- R. Pellegrin and C. Coates, "Executive and Supervisors: Contrasting Definitions of a Career Success," Administrative Science Quarterly, 1 (1957), 506-17.
- L. W. Porter, "A Study of Perceived Need Satisfaction in Bottom and Middle Management Jobs," Journal of Applied Psychology, 45 (February 1961), 1-10.
- L. W. Porter and E. E. Lawler, Managerial Attitudes and Performance, Homewood, IL: Irwin, 1968.
- L. W. Porter and V. F. Mitchell, "Comparative Study of Need Satisfaction in Military and Business Hierarchies," Journal of Applied Psychology, 51 (April 1967),
- R. D. Pritchard, "Equity Theory: A Review and Critique," Organizational Behavior and Human Performance, 4 (May 1969), 176-211.
- T. Roselius, "Consumer Rankings of Risk Reduction Methods," Journal of Marketing, 35 (January 1971), 56-61.
- M. J. Rosenberg, "Cognitive Structure and Attitudinal Affect," Journal of Abnormal and Social Psychology, 53 (November 1956), 367-72.
- J. B. Rotter, Social Learning and Clinical Psychology, Englewood Cliffs, NJ: Prentice-Hall, 1954.
- E. H. Schein, Organizational Psychology, Englewood Cliffs, NJ: Prentice-Hall, 1965.
- C. D. Schewe, "Selected Social Psychological Models for Analyzing Buyers," Journal of Marketing, 37 (July 1973), 31-9.
- J. N. Sheth, "A Psychological Model of Travel Mode Selection," Urbana, IL: Bureau of Economic and Business Research of the University of Illinois, Working Paper #291, November 1975.
- E. C. Tolman, Purposive Behavior in Animals and Men, New York: Appleton-Century-Crofts, 1932.
- W. F. Van Raaij, "Economic Psychology and Marketing," Oslo: Symposium "New Directions in Marketing," 1976.
- W. F. Van Raaij, Consumer Choice Behavior: An Information Processing Approach, Voorschoten: VAM, 1977.
- V. H. Vroom, Work and Motivation, New York: Wiley, 1964.
- M. A. Wahba and J. G. Bridwell, "Maslow Reconsidered: A Review of Research on the Need Hierarchy Theory," Organizational Behavior and Human Performance, 15 (April 1976), 212-40.
- E. Walster and G. W. Walster, "Equity and Social Justice," Journal of Social Issues, 31 (Summer 1975) 21-43.

INFORMATION ABOUT CONSUMER PROBLEMS BY THE POLICY-DELPHI PROCEDURE.
A NEW METHOD FOR EVALUATING THE MARKETING SYSTEM.

Solveig Wikstrom, Stockholm University

Abstract

The basic premise of this paper is that the marketing system does not perform adequately as seen from a consumer's perspective. For various different reasons the sellers' offerings do not sufficiently match the consumers' needs and wants. The performance of the socio-economic system to which the marketing system belong depends on adequate feedback. Existing information mechanisms are inadequate. In this paper a new method is presented for gathering information about the consumer situation for use as feedback. By this method a panel of informed persons assess the consumer situation in several rounds according to the so called Delphi procedure. By this method the gathering of subjective and objective information is made through one approach. Experience reported in this paper indicates that further development of the method should be carried out.

Introduction

In Sweden as in most industrialized countries intense activities are going on in the consumer policy area to strengthen the consumer's position on the market. Examples of important policy actions in Sweden in the 70's are:

1971: The Consumer Ombudsman, the market Court Act, the Marketing Practices Act, the Act Prohibiting Improper Contract Terms, the Door-to-Door Sales Act.

1972: The Food Office, the Food Act.

1973: The National Board for Consumer Policy, the Act on Products Hazardous to Man and Environment.

1974: The Consumer Sales Act.

1975: Decision about expanding the local consumer advisor activities.

1976: The New Marketing Practices Act.

There are more activities under consideration as a result of recent government commissions.

If we take these new laws and norms to be well motivated, they can be seen as indicators of deficiencies in the marketing system. Consumers as a whole or categories of consumers seem to have problems which consumer policy activities are supposed to counteract. What are the effects of these activities for the performance of the marketing system? What are the consequences for the consumers?

As a matter of fact there are no answers to such fundamental questions about the consumers' situation as:

- How often do consumers meet problems in their planning of purchases, buying and consumption of goods and services?
- What has been the trend over time?
- How serious are these problems?
- Are There product and service areas which have more problems than average?
- How much of the problems are the consumers themselves aware of?

In reality very little effort has been made to get this kind of feedback. To some extent this may be due to the lack of workable information mechanisms.

This paper presents some of the work carried out last year developing and implementing a new method for evaluating some aspects of the marketing system from

a consumer's perspective (Wikstrom, Eliasson 1976) This was done by having a panel of informed people assess the extent to which consumers meet problems and what implications these problems have for the consumers. The assessments were carried out in several rounds following the so called Delphi method.

In this case the Delphi assessment of the consumer situation was used as a method for evaluating what has been done up to date to improve the marketing system. I particularly want to emphasize that this new approach by no means fills the information gap pointed at above. It however does reveal new aspects of the problem not obtainable by existing methods and the experience gained indicates that further development of the method should be carried out. As a frame of reference for the presentation of the new approach a short overview is given of methods used at present.

Information generated by present methods

From the literature discussing criteria variables for evaluating the marketing system from a consumer perspective the following matrix can be arranged (Renoux, 1973, Lingoos, Pfaff, 1972, Olander, 1976, Andreasen, 1975)

FIGURE 1

Figure 1.

The information encompasses		
The databasis are ...	The micro marketing system	The macro marketing system
<u>Subjective</u>	- consumer complaints - consumer satisfaction/dissatisfaction surveys	Studies of consumer attitudes to the business policy in advertising, product development, product safety etc.
<u>Objective</u>	Studies of - number and location of retail outlets, shopping distances - trends in life span, quality, prices and repair costs of goods - the occurrence of unsafe and unhealthy products - consumer shopping pattern	Studies of - innovation capacity of the business - restraints in competition - advertising volume - brand differentiation etc.

Subjective data basis means statements based on individual consumers' evaluations of their situation and the state of their environment. Objective data basis implies that data used are not directly dependent on the consumers' evaluations of the situation. Information about the micro marketing system includes aspects

such as individual products and firms, shopping possibilities, the buyer situation in the outlets, the consumption and repair of products etc. The macro marketing system covers the total system. The evaluation concerns its functions from a wholistic view.

The advantage of information based on subjective indicators is that these reveal what the consumers themselves perceive as their problems and as deficiencies in their marketing environment. These types of indicators can not, however, be used as a decision basis either by public policy regulators or by business planners without analysis and evaluation. One has to clarify what, if any, are the alternatives available within the present technology and the resulting costs to consumers. The consumer should be informed as to whether or not there are alternatives. There are serious "consumer problems" which can not be dealt with within the marketing system. In this case the government must step in and provide some solution outside the marketing system, e.g., in public transportation, in housing, or in education.

The main weaknesses of the subjective indicators are their dependence on consumer aspiration levels and their inability to generate information about those problems which the consumers may not be conscious of.

Evaluations based on objective indicators, on the contrary, have the advantage of recognizing problems which the consumers are unaware of. Trends in this kind of data can also be used as indicators for future problems.

There are however validity problems in using objective indicators. You really do not know how close a relationship there is between the indicators and the state you want to measure. Besides that you do not know if the set of indicators used is sufficient for measuring a certain consumer state.

The conclusion is that either subjective or objective indicators separately are sufficient as information instruments. Both types of indicators are needed, which implies that the information gathering becomes extensive and costly. Besides that there are still unfilled gaps.

The new method of evaluating the consumer situation by using assessments by a panel of "experts" combines the subjective and objective approach. By treating the gathering of subjective and objective information through one approach we are able to put them on a relative or comparative basis in terms of what we believe is now an efficient process for producing relevant information. First the panels' professional background implies good insights into what is called objective indicators. Second these people are also well informed and knowledgeable consumers themselves. Therefore, they should have qualifications to integrate subjective and objective aspects in their assessments.

The panel assessment in the actual method study has been directed towards micro aspects, but in a systematic way so that the marketing system in large has been covered. This way micro and macro aspects have been integrated.

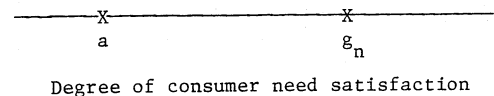
Referring to the matrix in Figure 1, the position of the Delphi assessment is between subjective and objective and between micro and macro and thus integrating these aspects.

Definition of the Consumer Problem Concept

In Russell Ackoffs terms (1962) the concept problem implies a tension between what a person should achieve and in fact does achieve, i.e., a deficiency in goal fulfilment. What a person should achieve has however

to be kept within what is technologically possible and attainable. With the restriction to what is attainable the utopias are out of question. This definition is illustrated in Figure 2.

FIGURE 2 Consumer problem as a deviation between the achieved state (a) and the desirable and attainable state (g_n).

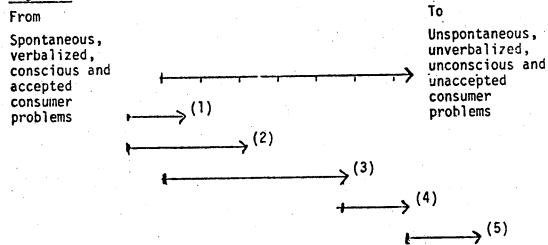


Defining consumer problems in this manner presupposes knowledge about consumers' achieved states, knowledge about attainable states and values about what is desirable. Differences in these components lead to differing views about how widespread and urgent the consumer problems are.

When defining consumer problems it is possible to distinguish between a number of concepts related to this problem concept in the way indicated in Figure 3.

- (1) Problems from consumer complaints. Spontaneous dissatisfaction measured by number of consumer complaints.
 - (2) Problems from surveys. All problems which the consumers are aware of and admit. This concept is used in surveys of consumer satisfaction/dissatisfaction.
 - (3) Unconscious problems as well. Relevant subsets of problem areas (1) and (2) as well as some specific problems not admitted or perceived by the consumers. We know that in some situations the consumer denies his problems. He rationalizes his poor decisions in order to reach internal harmony, acknowledging information which supports his decision and avoiding stimuli which are in conflict with it. These mechanisms are explained by the theory of dissonance (see e.g. Festinger, 1964). The consumer would not admit these kind of problems.
- Then there are problems which the consumer is unaware of for another reason. He lacks information about his possibilities to attain better states. Better alternatives already exist, but the consumer does not know about them.
- (4) Problems caused by lack of dynamics. Better Alternatives are attainable with present technology including better organization of the marketing systems. Because of lack of dynamics these alternatives are not realized.
 - (5) Problems from deliberate behavior. Certain kinds of consumer behavior give rise to a state which, according to established norms, is poor. An inadequate diet and consumption of alcohol and drugs are illustrative examples. Buying groceries very frequently is another kind of behavior which consumer policy planners regard as a problem. The consumers themselves however behave as they are used to. According to their own present values their state apparently is better than the other state which the planners consider highly preferable.

Figure 3 The consumer problem concept and subconcepts



In Figure 3 the consumer concept is seen as a vector. The subconcepts - the numbers refer to the explanatory text above - are arranged along the vector after degree of verbalization, consciousness etc.

It is to be noticed that all consumer dissatisfaction can not be regarded as consumer problems. If there are no better alternatives attainable, the consumer dissatisfaction does not indicate consumer problems, according to the definition used. This is illustrated in Figure 3. Parts of the subconcept (1) in (2) are located outside the consumer problem vector. Seen from an individual consumer perspective the type (4) problems are not consumer problems either. The better state is not desirable for the consumer.

It is also to be noticed that we do not know the relations between the subconcepts. The size of these vectors in Figure 3 are arbitrary.

All the above problem areas have been discussed in the literature; however, the distinction between the problem types is hardly ever made explicit. When policy actions are considered these distinctions should be noticed.

Our exercise was aimed at clarifying problems of type (3) and (4). Using a panel with knowledge about consumer behavior as well as present and potential supply gives an opportunity to take into account those problems which individual consumers are unaware of or are unwilling to admit.

In the operationalization we tried to exclude (4) thus measuring the amount of consumer problems perceived by fully informed but not necessarily rational consumers in a narrow sense. The working hypothesis was that people strive at using their economic and other resources in such a way that they get as good a life as possible. In this goal seeking behavior conflicts arise. In balancing these conflicts people sometimes consciously choose a poor alternative in fulfilling a certain sub-goal. In a narrow perspective the decision seems irrational, from a wider perspective quite rational.

Design and Implementation of the Panel Assessment

The panel composition

In selecting the panel members two qualities were important. First, the members should have good knowledge about consumer behavior and about present and potential supply, i.e., about data earlier described as objective indicators. The knowledge areas aimed at were identified from the members' employment or professional position. Second, as has been emphasized above, values hold an important role in defining consumer problems. Considering this, efforts were made to include members

viewing the consumers' situations from different perspectives and representing different value structures. As a consequence the panel was composed of the following categories.

- (1) Local consumer advisors (LA)
- (2) Representatives of public authorities (PA)
- (3) Representatives of the business sector (BR)

The guidelines for the panel size was about 30 participants, which is considered sufficient for a Delphi-procedure. The final panel consisted of 9 from the (LA) category and 10 each from both others.

Method for Data Collection

Data were generated by a survey in different rounds by means of the so called Delphi method. With this method the participants give their answers individually on written questionnaires. The individual answers remain anonymous. The answers of the panel are reported back to the panel in statistical form as median and quartile answers. Questions and remarks are also fed back to the panel, together with explanations and clarifications from the panel director. On the basis of such information the participants reassess their previous statements in new rounds.

The Delphi inquiry utilizes well-informed people's subjective judgements. The method was originally developed as a forecasting tool. It has however been used more and more to generate approximations and judgements about other situations where "hard data" are difficult to obtain. The method has been used for instance to define different opinions of important social issues (see eg. Jolson, 1975, and Dahl, 1974), as a basis for decision making in marketing (Jolson and Rossow, 1971) and as a general tool in evaluation research. For more details about the method, see eg. Linstone, Turoff (1975) or Wikstrom (1973).

The Questioning

The panel assessment concerned two different aspects:

- a) how frequent and how serious are consumer problems in planning of purchases, buying and consumption of goods and services.
- b) what desirable and feasible policy actions are to be taken to counteract these problems.

The latter aspect was included not only to give a wide spectrum of ideas about possible action, but to highlight what ideas there are about causes to these problems.

The panel was provided with some information about the general idea behind this assessment and a presentation of the outline of the study. This was done by the conceptual model in Figure 4.

FIGURE 4

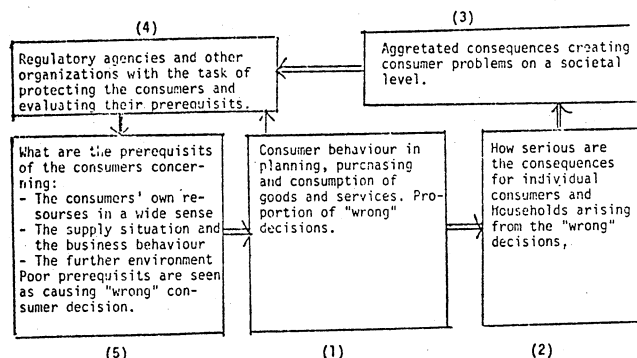


Figure 4

The first part of the assessment considered (1) and (2) in Figure 4. In this the panel was asked to use available data about (5). The outcome of this assessment gives (3). In the second part, (4) and (5) are dealt with. The panel was asked to suggest possible and feasible actions to counteract the stated problems. The panel was exhorted to seek information and ask questions if anything seemed confusing.

The Operationalization of the Consumer Problem Concept

Consumer problems were seen as situations in the planning, buying and consumption process, where consumers unwittingly make "incorrect" decisions, i.e., behave in a way which affects their budget or need satisfaction negatively. In order for an action to be defined as a wrong decision, it should have clearly noticeable negative effects. "Unwittingly" means that the consumer would not have behaved in that way if he had foreseen the consequences and if there had been other alternatives at hand.

The frequency of problems was measured as the average number of wrong decisions the consumer makes per 100 decisions. How serious the wrong decisions were was measured according to a 3 point ordinal scale.

As a second part of the study the panel suggested steps to be taken to counteract the problems they had established. In the second round the panel evaluated all measures suggested.

To get a systematic view of the consumer situation the panel was asked to assess the amount of problems in the following situations. For durable goods and more expensive services such as cars, private homes, boats, dishwashers, television sets and package vacation tours, the question was how often consumers make "incorrect" decisions on an average when they:

1. plan and consider and ultimately decide to buy a certain product or service.
2. choose product class, size, price level, and retail outlet.
3. carry out the buying: choose how to pay, agree to purchase conditions, make complaints etc.
4. use the product or service.
5. buy repair services.

For convenience goods the incorrect decisions concern:

6. the shopping pattern including frequency of shopping choice of retail outlets, time, transportation etc.
7. the choice of brand, package, size, price level, etc.

These steps in the buying process are based on ideas from Gredals typology (Gredal, 1971).

The panel was also to mention and assess consumer categories and product areas with particularly high frequency of wrong decisions.

The Delphi process

The written questioning consisting of two rounds was carried out in spring 1976. The panel answered each questionnaire within 3 weeks. After the first round some questions had to be rewritten and new instructions given. Misunderstandings of the concept and the assessment procedure had appeared.

After a working report had been finished and sent to the panel members a telephone round was carried out in early 1977. The subject was the panel's experience of the assessment process and the opinions about the final results.

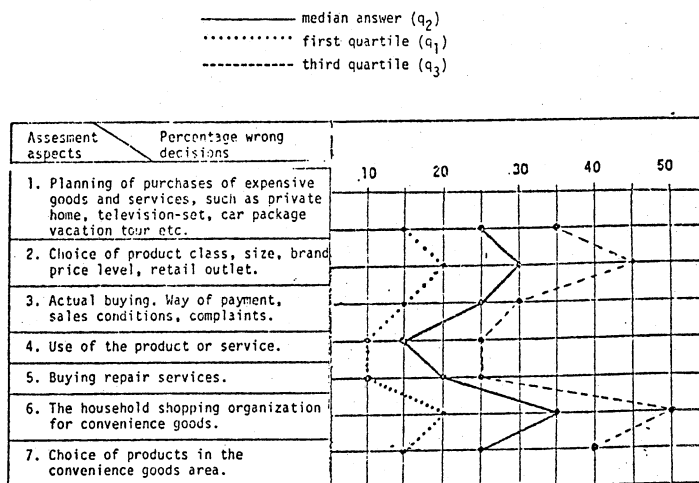
Some Results

The Amount and Seriousness of Consumer Problems

The assessment of the frequency of problems for the average consumer is shown in the profile below.

FIGURE 5

The assessment of the percentage of total number of wrong decisions made by the average consumer in planning purchases, buying and consumption of goods and services.



According to the median assessment there are problems for the consumers in 15-35 percent of the decisions in the different steps in the buying and consumption process.

The frequency of problems is one dimension, the seriousness of the problems is another.

In analyzing the panel assessments a total score was used for the negative economic and need satisfaction consequences. As mentioned before the consequences were measured according to a 3 point scale. The most serious consequences were given 3 points, the next group 2 points and the least serious 1 point. As shown by Figure 6, the first two aspects, planning in a wide sense, has the highest scores for negative consequences. As a whole negative economic effects are bigger than the need satisfaction effects. For the choice of convenience goods the opposite holds good.

FIGURE 6

Importance scores for the negative economic and need satisfaction consequences of consumer's wrong decisions.

The diagram shows that the (LA) representatives tend to consider the consumer situation to be more problematic than the other categories. The differences are, however, not as clear as I had expected.

For two reasons I had anticipated that the (BR) category would have given the most positive view of the consumer situation. First, business people regard consumers as a powerful group. The consumers "vote" by buying or not buying. From the business people's view they are not very easy to convince. Second, for tactical reasons this category could be expected to be optimistic. A problematic consumer situation calls for more regulations, which are mostly directed against the business.

As the diagram shows, there are no clear differences between the (BR) and (PA) categories. The assessments by these categories by no means reflect the differing views between the same categories in the public debate. After having analyzed the assessments and discussed the results with the panel in the telephone round the small differences between the categories can be explained. In the assessment process individual values and personal experience of buying have had a much stronger impact than group values and professional experience. In other parts of the assessment, concerning possible policy actions to be taken to counteract the stated problems, there was much more explicit groups differences.

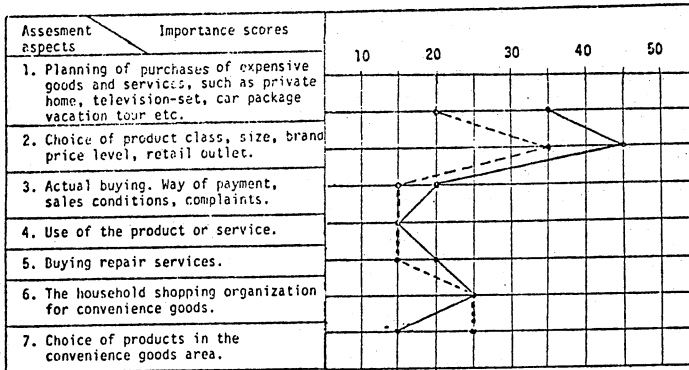
Suggested policy actions

The panel as a whole suggested about 60 policy actions of very different character. About half of them were business oriented, one third were market oriented and the rest were consumer oriented. The panel evaluated these actions remarkably positively according to desirability and feasibility.

In order to give an idea of what kind of actions the panel suggested, some examples are quoted. All the three panel categories rated these actions as definitely feasible and desirable. There is also a set of more radical actions suggested, about which there were disagreements.

- Actions to improve the consumers' planning of purchases:
- Improved education of young people and adults in consumer issues, particularly in budgeting.
 - Increased consumer information in present and new forms.
 - Extensive use of the television for consumer information.
 - Expanding the local consumer advisor and consumer information activities.
 - Increased product information based on product tests from the National Board for Consumer Policies.
- Actions to improve the buying decisions:
- Compulsory use of simple contract forms.
 - Compulsory information about actual interest in all installment buying.
 - Develop standard forms for consumer complaints.
- Actions to improve the use of products:
- Claim for instructions easy to understand for normal use and when troubles occur.
 - More consideration to an easy use in product design and product development.
 - Seller responsibility for complete and correct information about product use.
- Actions to improve the household shopping organization:
- Prevent the closing down of retail outlets in rural areas.
 - Increased consumer influence in local retail planning.
 - Prohibit further hypermarkets.
- Actions to improve the choice of products in the convenience goods area.
- Compulsory price comparison.

Economic —————
Need satisfaction - - - - -



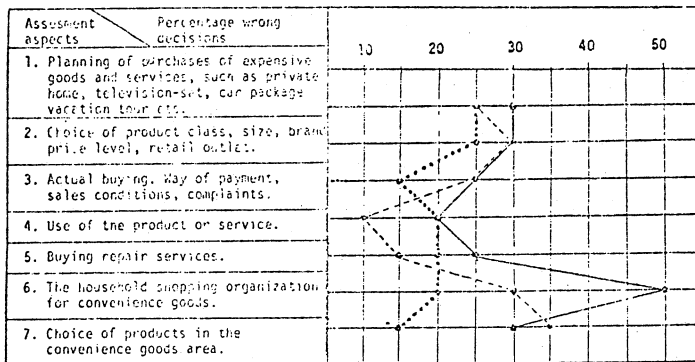
x) As mentioned before the consequences were measured according to a 3 point scale. The most serious consequences were given 3 points, the next group 2 points and the least serious 1 point.

An important part of this study was to inquire into the differences in values and views between the three subgroups of the panel. These differences are shown in Figure 7.

FIGURE 7

The assessment of the percentage of total number of wrong decisions according to the median assessment of the three subpanels.

————— Local consumer advisers (LA)
- - - - - Representatives of the public authorities (PA)
..... Representatives of the business sector (BR)



- Educate the consumers to read the declarations of content.
- Standard packs.

The business representatives however opposed some actions directed towards their own sector, such as:

- Increased consumer insight into and influence on the product development.
- Reducing advertising volume.
- Brand differentiation prohibited.
- Product variety reduced in certain areas.
- Reduced "bargain" prices.

The professional background of the panel members seemed to have a stronger impact in the assessment of possible policy actions than in their stating the amount of consumer problems.

Analysis and Conclusions

What do these panel assessments measure and how valid are the results? Compared with the subjective and objective indicators discussed in an introductory passage the panel assessment aims at measuring the total consumer problem concept. This measurement procedure covers the problems which the consumers are aware of as well as those they do not know about.

The subjective and objective indicators illustrated in Figure 1 are fairly accurate measures. They are, however, partial and we do not know how valid they are as measurements of the total consumer problem situation. This issue was discussed in connection with the different subconcept in Figure 3.

Validity includes, however, good reliability. In a panel assessment we know very little about the measurement procedure, about what model the individual members use and about what data they feed into the model when they do their assessments. We know however that their assessments are some sort of syntheses of a very complex reality. In a Delphi procedure using panel assessment as the data source, the reliability problem is a validity problem on the individual level. Do the individual members process the data they are supposed to (see Figure 4) and do they generate good information? We just do not know. With these questions one enters another theory of science. Keywords such as dialectical and phenomenological indicate what kind of theory is referred to. According to this theory the reality is too complex to apply any concrete validity criteria. One criterion suggested is to what extent the actors involved accept the results and the interpretations. This subject is however very difficult and more thinking is needed.

The quality of the information gathered was also discussed with the panel. Their reaction was that the assessments could be used as rough measures. I would say that the outcome in this case should be expressed by the interquartile rather than by the median figures (see Figure 5). This means that you can not from the figures say if the proportion of "wrong" decisions is 15 or 35%, you can however see that some steps are more problematic than others.

Another vital question is how the stated problem situation should be interpreted. Do the figures indicate a satisfactory consumer situation or not? According to the panel the consumer situation is unacceptable. Policy actions should be taken. The panel has suggested a long list of possible actions, most of them directed toward suppliers. You can also see from the results to what extent different panel members agree on the suggested actions.

The policy Delphi procedure is not only a possible information and evaluation instrument but is also a tool

for generating ideas about how the stated problems should be cured.

In our study the Delphi assessment was used to gather information about the effects of the policy actions taken up to now on the consumer situation and to evaluate that situation. The results indicate that there are considerable problems. For this kind of statement there is no need for very accurate measurements. The results also show comparisons between different areas. For this kind of assessment there is a need for a somewhat higher accuracy. I think that the method is valid for this kind of assessment too. One has however to take into consideration that if a special aspect has been openly debated recently this may effect the results considerably.

In comparing of situations at two different points in time a rather high accuracy in the measurements is needed. The same holds true if the aim is to compare the situation between different countries. At present I would be hesitant to use the method for this kind of comparisons. More knowledge, experience and thinking are needed.

Finally when judging the value and the possibilities of this method one should keep in mind that this study was a test of the method. It is possible to improve the results considerably by a better design of the assessment procedure.

You should also, when judging the value of this method, ask how good results you could get for the same money using another method. In this aspect I think the panel assessment is quite competitive.

References

- R.L. Ackoff, Scientific method. Optimizing applied research decisions. New York. (1962).
- A.R. Andreasen, Consumer dissatisfaction and market performance. Paper presented at Fourth International research seminar in Marketing, Senanque Abbey, Gordes, France. May (1977).
- A.W. Dahl, Delphic and Interactive Committee Processes in a Comprehensive Health Planning Advisory Council: A comparative Case Study, PHD Dissertation, The Johns Hopkins University. (1974).
- G. Eliasson, Problemformulering och problemuppfattning i den konsumentpolitiska planeringen. Arbetspapper. Foretagsekonomiska inst. Stockholms Universitet. (1976).
- K. Gredal, Moderne forbrugeres motiver og adferd (The motives and behaviour of modern consumers), Kopenhagen. (1971).
- L. Festinger, Conflict decision and dissonance. Stanford, California. (1964)
- I.A. Jillson, The National Drug-Abuse Policy Delphi, in Linstone & Turoff, The Delphi Method. (1975)
- M.A. Jobson and G.L. Rosson, The Delphi Process in Marketing Decision Making, Journal of Marketing Research. November (1971)
- J.S. Lingoos and M. Pfaff, The Index of Consumer Satisfaction: Methodology. In M. Venkatesan (ed), Proceedings of the Third Annual Conference of the Association for Consumer Research. (1972)

H. Linstone and M. Turoff (ed), The Delphi Method. Techniques and applications. Addison-Wesley. (1975)

Y. Renoux, Consumer Dissatisfaction and Public Policy. In F.C. Allvine (ed), Public Policy and Marketing Practices, Chicago: American Marketing Association. (1973)

S. Wikstrom, Vara dagligvarors distribution. En framtidsstudie. (The Food Marketing Process: A Future Study), Bonniers. (1973)

S. Wikstrom and G. Eliasson, Konsumenters problem vid planering, kop och konsumtion av varor och tjanster - En policy-delphi studie. (An assessment of consumer problems in planning purchases, buying and consumption. A policy-delphi study). The Department of Business Administration Stockholm University. Working paper. (1976)

F. Olander, Consumer satisfaction - a skeptic's view. Paper presented at NSF Workshop on Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction, Chicago, April 11-12, 1976. (1976)

THE PERCEPTION OF FOREIGN PRODUCTS IN FRANCE

Gary Baumgartner, Maitre de Conférences
Alain Jolibert, Maitre Assistant

Abstract

Despite numerous articles which have been published, there seems to have been no real advancement in the development of theories which might explain why consumers view foreign products differently than they do domestic products. This paper reports a study of French perception of foreign products using the theory of perceived risk as criteria for the choice of products to be studied.

Studies of Foreign Product Perception

To date the published research in the field of foreign product perception can be classified into two types: research which essentially eliminates the product from consideration, and research on the perception of apparently randomly chosen specific foreign products and their images.

The first type can be classified as national product stereotype research. The theory has been that attitudes specific to a country emerge when products from that country are perceived. However, these studies are based on either abstract notions such as "Italian Products", or else on products with such low profiles that the consumers naturally only perceive the national origins of the product. For example, Schooler (1965) in studying the Central American Common Market found that evaluations of beige fabric and fruit juice varied according to the country of origin; these evaluations were related to preconceptions regarding attitudes toward the people of that country. Reiersen (1966, 1967) in studies involving American consumers evaluating "Italian Products" and "Japanese Products" found that foreign product quality tended to be stereotyped and that the communication medium used to promote the particular foreign product influenced the stereotyped image. Nagashima (1970) found that when Japanese and American consumers evaluated the notion of "all products" made in a particular country, attitudes varied according to the opinion of the country of origin. Anderson and Cunningham (1972) in using the general classification of "foreign products" found significant differences in the socio-demographic and psychological characteristics of those consumers who were favorably disposed toward foreign products and those who were not.

Whereas these types of studies permit a certain generalization of their results, such generalizations tend to be inappropriate or misleading when applied to specific products. In fact, this is the rationale adopted for the second type of research which involves very specific products such as toys, autos, cameras, etc. Unfortunately, the fact that such products are chosen in the absence of a theoretical framework eliminates the possibility of generalizing the results. A good example of this type of study is presented by Etzel and Walker (1974) who present evidence to show that the general use of national product stereotype can be very misleading. However, the alternative they offer is to restudy each individual product.

Perceived Risk and the Perception of Foreign Products

One theory in consumer behavior which seems particularly appropriate to the study of foreign product perception

is the theory of perceived risk. Perceived risk related to purchase depends upon two determinants (Cox, 1967). First is the degree of uncertainty concerning the suitability of the product. Second is the importance which the consumer affects to the possible adverse consequences of the purchase. Both of these determinants could be intimately linked to foreign products. In addition to the uncertainty linked to the product proper, there is an additional uncertainty due to the origin of the product. As to the possible adverse consequences of the purchase, Jacoby and Kaplan (1972) have operationally defined and studied 5 different varieties of perceived risk, each linked to some form of adverse consequence of the purchase. These five components of perceived risk are financial risk, performance risk, physical risk, psychological risk, and social risk. It would seem that the consumer's stereotype of the product's foreign origin could in fact influence his appreciation of the particular risk involved in the purchase of the product. For example, it would seem that the purchase of life insurance from a company based in a country which is actually experiencing economic difficulties would increase the risk perception involved in the purchase.

The working hypothesis for the study was that perceived risk is a significant component in the judging of foreign products. It was thus expected that:

- foreign products with different types of risk components would be judged differently,
- foreign products would be judged differently than domestic products involving the same risk.
- the perception of the risk component of the product is conditioned by the national origin of the product.

Method

A questionnaire was developed to obtain a measure of consumer interest for each of 16 products varying according to 4 different types of risk and with 4 different national origins. The data were analysed by analysis of variance (ANOVA) techniques.

The Product Classes

Four product classes were chosen to represent one or several specific components of perceived risk:

- playing cards (absence of perceived risk)
- life insurance (financial and performance risk)
- cough syrup (physical risk)
- winter coat (social and psychological risk)

We started with the list of products published by Jacoby and Kaplan (1972) from which we retained certain products and eliminated or substituted others. Certain ones were eliminated on the basis of too strong a domination of the French market by a particular country. For instance, color television was eliminated because German made televisions are quite popular in France whereas English made televisions are unknown as such. Vitamins and aspirins were eliminated as physical risk products on the grounds that the French consumption pattern of

vitamins and aspirins and consequently their perception of these products are very different from that of Americans. Cough syrup is a product which in the French context closely approximates the American perception of vitamins and aspirins. Products were chosen to be appropriate for both sexes.

The National Origins

In addition to French, which was included as a control and as a basis of comparison, the foreign origins retained for study were American, English, and German. These were chosen according to the criteria of equivalence and relevance. The study was limited to the comparison of countries considered to be equal, respected trading partners of France. Countries with low prestige profiles due either to low economic development or to an unfavorable foreign product stereotype were eliminated. This was done on the basis that an overall prejudicial image could dominate so strongly that the perception of the differential product risk, might be affected.

Questionnaire

A seven point scale was constructed with scale values ranging from "extremely interesting" (1) to "not at all interesting" (7). The respondents were placed in a purchasing situation and asked to evaluate each one of 16 products (4x4) presented in random order. Sociodemographic as well as political data were collected but were not analyzed in this study.

Subjects

120 French consumers were interviewed of whom 108 complete questionnaires were retained for analysis. Characteristics of the sample indicate a slight predominance of female (57%) over male respondents. A slight bias of educational level was present with 39% of the respondents had attended at least one year of undergraduate study.

Data Analysis

Given the nature of the data, a fixed-effects 4 x 4 repeated measures randomized block factorial design was selected as the most appropriate analytical technique (Kirk, 1968). This design is justified since each of the 108 subjects was exposed to all levels of the two factors (4 nationalities x 4 product risk classes). Because of this, each subject is technically considered a block; this means that there are 1728 (4 x 4 x 108) cells in the design and no subject x factor interaction is possible.

Analysis and Results

The model for a repeated-measures analysis is appropriate if the variance-covariance matrices are equal and if the pooled variance-covariance matrix is symmetrical (Kirk, 1968). The homogeneity of variance was tested using Hartley's F_{max} test. At a risk of $p < .05$, F_{max} was not significant and thus the assumption of homogeneity of variances within cells can be assumed.

The ANOVA results are presented in Table 1. All main effects as well as the interaction effects are significant ($p < .05$).

TABLE 1

Analysis of Variance for Perceived Interest

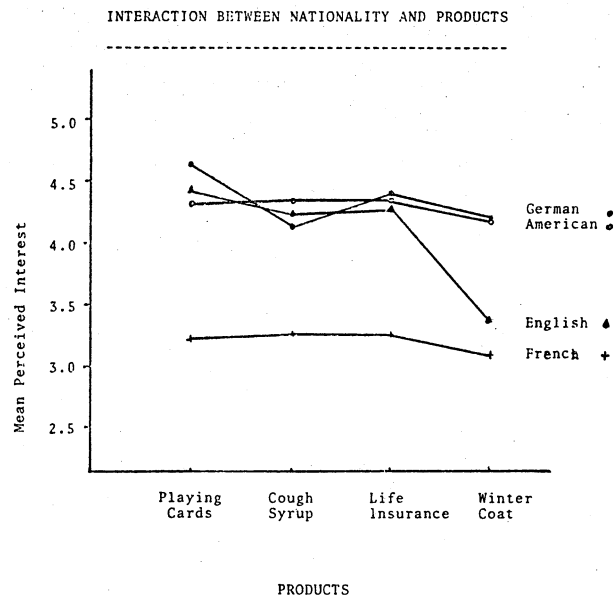
SOURCE	df	MS	F
Subjects	107	6.52	
Nationality (N)	3	163.76	109.60 *
Products (P)	3	27.47	18.38 *
NxP	9	5.11	3.42 *
Error	1605	1.49	

* $p < .05$

The study of each effect as well as the interactions were analyzed by a multiple comparison of the means. Dundan's Multiple-Range test, which is a more stringent test procedure than the t test, was used. The significant results are summarized below.

1. French consumers prefer French products to foreign products.
2. All products were judged equal with the exception of the winter-coat which was judged more interesting than the others.
3. Figure 1 presents the interaction between nationality and product. The most preferred combination was the French winter coat, and German playing cards the least preferred combination. A notable exception to the French perception of foreign products is that the English winter coat is not judged as unfavorably as other foreign products. However, the French winter coat was still preferred.

FIGURE 1



Discussion

The results indicate that the French consumer has a very strong preference for domestic products. This is not surprising in light of the values inherent in the French culture, notably individualism and nationalism (Dubois, 1972). This means that the manufacturers of foreign products should not in general stress their origin in brand image strategy.

Perhaps the most interesting results of the study is the extremely good showing of the English winter coat. This is all the more surprising because French winter coats have an excellent reputation, even outside of France. Since the winter coat assumedly implies an inherent social and psychological risk, we advance the hypothesis that foreign products in France are more likely to be purchased if they fall into that category. Indeed this could explain why certain foreign products are popular in France.

The English winter coat might have been appreciated by the French consumers because the English presently enjoy a favorable national stereotype in terms of fashion. From these observations, we advance the second hypothesis that for the French consumer, favorable national stereotypes are ineffectual unless they are related to products with a certain degree of social and/or psychological risk. The implications of this hypothesis for foreign manufacturers would be the following. If a product enjoys a favorable national stereotype in France, the product's publicity could perhaps best be oriented towards an emphasis on the social-psychological risks of the product class. This seems to be the area where foreign products compete most effectively with French domestic products.

Naturally, the results of this study must be tempered in the light of the relatively small sample size as well as the characteristics of the sample. Nevertheless, the results indicate that the theory of risk perception provides an interesting insight on the judgment of foreign products. More important perhaps is that specific objective criteria founded on a theoretical framework were for choosing the products used in this study. Because of this, generalization of the results become possible. In our case, we limited these generalizations to the formulation of hypotheses because of the study's restricted nature and its small sample size. Nevertheless, we hope that our study will act as a stimulus in the area of foreign product perception. Indeed, much remains to be done in the way of practical research as well as theoretical development before foreign product perception is understood.

References

William Anderson and William Cunningham, "Gauging Foreign Product Promotion," Journal of Advertising Research, Vol. 12 (February 1972), 29-34.

Donald F. Cox, Risk-taking and Information Handling in Consumer Behavior, Boston: Harvard University Press, 1967.

Bernard, Dubois, "A Cultural Approach to the Study of Diffusion and Adoption of Innovations," Proceedings of the Third Annual Conference of the ACR, 1972, 840-50.

Michael Etzel and Bruce Walker, "Advertising Strategy for Foreign Products," Journal of Advertising Research, Vol. 14 (June 1974), 41-4.

Jacob Jacoby and Leon Kaplan, "The Components of Perceived Risk," Proceedings of the Third Annual Conference of the ACR, 1972, 382-93.

Roger E. Kirk, Experimental Design: Procedures for the Behavioral Sciences, Belmont, Calif.: Brooks Cole, 1968.

Akira Nagashima, "A Comparison of Japanese and U.S. Attitudes Toward Foreign Products," Journal of Marketing, Vol. 34 (January 1970), 68-74.

Curtis Reiersen, "Attitude Change Toward Foreign Products," Journal of Marketing Research, Vol. 4 (November 1967), 385-7.

Curtis Reiersen, "Are Foreign Products Seen as Traditional Stereotypes," Journal of Retailing, Vol. 42 (Fall 1966), 33-40.

Robert Schooler, "Product Bias in the Central American Common Market," Journal of Marketing Research, Vol. 2 (November 1965), 394-97.

EXPOSURE TO CULTURAL ACTIVITIES
AND OPINION LEADERSHIP

E. Langeard, University of California, Berkeley
M. Crousillat, Université d'Aix-Marseille
R. Weisz, Université d'Aix-Marseille

Abstract

Consumer services have recently been subjected to systematic research at both conceptual and empirical levels (4,8). This paper focuses on a service activity which objects very much to the use of marketing terminology and concepts and even more so when it is located in the European environment. However, cultural activities and, among them, performing arts have serious attendance problems and difficult communication channels with both their actual and potential audience. Several studies have already been made of attendants' behavior, taking into account usual socio-demographic characteristics (5,6,9).

The objective of this paper is to study the phenomenon related to interpersonal communication within groups of consumers more or less exposed to cultural activities. The first part is devoted to an analysis of the communication process which is precominant for cultural or artistic services. The second part presents the research methodology and the set of hypotheses, and the statistical findings of the survey made the third part. Finally managerial implications that might be drawn from the survey are given.

Interpersonal Communication
And Attendance Of Cultural Services

The decrease of attendance has been a major concern for the last ten or fifteen years in many countries. Several surveys of live theaters' and movies theaters' attendance have been made. From them it is known that in 1974 63% of the total French population have been at least once a year at a movie theater and 69% of the U.S. population of 12 major cities (12, 14). Also in 1974, 22% of the total French population have seen at least one amateur or professional theater play and 31% of the U.S. population of 12 major cities (12, 14). The data do not provide us with information only on attendance levels but also give us clues about why attendance levels are so low. Take the example of theater attendance. Pricing is not a major issue, at least in France where tickets from subsidized theaters are not much more expensive than movie theater tickets. Theater attendance is a highly socialized process. Willingness to go to theaters is related to education level and social status. Translating willingness to go into actual attendance requires the support of a group providing key-information, companionship, and a discussion platform. When people who do not go anymore to theaters are asked why, the reason they give most often is they have been cut off from the group with whom they attended or their new companions do not like to go to theaters. Sixty percent of the population which is not exposed to theater performance give as a main explanation the lack of guides, of "persons who know" among relatives and friends (12). Interpersonal communication has a strong influence on theater attendance and the lack of funds for mass communication which generally prevails increases even more this influence. Mass communication media are more widely used in other performing arts such as film. Reading as a cultural activity is more and more under the pressure of huge advertising campaigns. In both cases, word of mouth has still devastating or

highly positive influence. Therefore, it is necessary to have a better understanding of the consumer search and communication process in the area of cultural services.

From prior research, a conceptual framework has been built which can be applied to many different types of service situations (3). Facing services, consumers have a problem of confidence. It's difficult to test or try services without the occurrence of an excessive cost. Without experience of the service and because its image is unclear, consumers do not know if it fits their needs or not. Finally, the performance and the quality of many services vary unexpectedly. These three remarks apply to cultural services. Potential customers try to sustain their confidence through an active search process of information from those who know, those who have experience the services. It creates optimal conditions for the existence of personal influence.

Methodology and Hypotheses

The self-designated method has been chosen for this survey, with the use of three questions based on Katz and Lazarsfeld (7) and Rogers and Cartano (11) researches. Identifying opinion-leaders by self-report has been validated (2,10,13) although some accuracy is sacrificed for economy and expediency (1).

In a large city of the south of France, three consumers' groups have been investigated in relation with the most important Repertoire Theater Company performing in the city on a permanent basis.

The first consumer group is made of subscribers to yearly programmed events. They agree in advance to attend a minimum number of shows at specific dates. They are given a fifty percent discount on ticket prices.

The second consumer group is made of non-subscribers, paying full price for a show they chose to attend.

The third consumer group is made of a sample of the general population, with quotas based on the repartition of the subscribers' group among occupational categories.

The questionnaire has been sent by mail to the subscribers. A stratified sample of 1,200 people has been used out of 6,000 subscribers. 529 answered the questionnaire which gives a 45% response rate. The questionnaire has been administered by interviews to 202 non-subscribers when attending a show. The sampling method introduces a bias and the probability was high to get a bigger share of "heavy-users" among non-subscribers. The general population sample has been interviewed at home. 205 persons answered the questionnaire, half male, half female. In the two other samples females represent 60% of the total, and the males 40% only.

Theater is generally considered as the most accessible of the performing arts outside of the mass media (compared with opera, ballet, symphony). It was decided to compare exposure to theater with other kinds of cultural activities. We selected one which is directly

related to mass communication: films in movie theaters, another one which does not require to go out and attend to a specific performance, which is book reading.

The questionnaire has been designed in such a way that provides information about universal leaders versus specialized leaders for theaters, films, and readings, and the overlaps between them. The relationship between the intensity of exposure to a specific cultural activity and the amount of information received and diffused by the people has also been investigated as well as the communication process being used. A set of hypotheses has been formulated in accordance with the primary objectives of the research.

On one hand the research had to verify if characteristics generally associated with opinion leadership would show up when studying cultural opinion leaders (H₅, H₆, H₇); on the other hand it had to explore any relationship between opinion leadership propensity and the willingness to subscribe and to commit oneself to a whole program of cultural events (H₂). Besides these two phenomena a third one was added which is the relationship between the extent of overall cultural activities someone has in the three domains (theater, films, reading) and his categorization as an opinion leader and as a theater subscriber (H₁ and H₃). Last it was expected that opinion leadership overlap would be verified between the three categories of cultural activities (H₄). Therefore the set of hypotheses is the following:

- H₁ Subscribers are not only attending more theater performances than non-subscribers; they are also seeing more films and reading more books.
- H₂ Subscribers have a greater propensity to be opinion leaders.
- H₃ Cultural opinion leaders are more exposed to cultural events or more committed to cultural activities than non-leaders.
- H₄ Opinion leadership overlaps exist between the three sub-categories of cultural activities taken into account in the research survey.
- H₅ Opinion leaders look more for information in the mass media than non-opinion leaders.
- H₆ The influence of personal sources of information is stronger on non-opinion leaders than leaders.
- H₇ Cultural opinion leaders are disseminating more cultural information than non-opinion leaders.

Findings

All hypotheses have been confirmed totally or partially except the one dealing with opinion leadership overlap and the one expecting theatre subscribers to be "heavy users" of performing arts. One of the unexpected findings is the lack of openness of the subscribers' group. This has important marketing implications. This study has demonstrated once more, and after many other researches with similar results, that it did not exist a significant relationship between opinion leadership and socio-demographic variables such as age, sex, occupation, family status or income.

1. Subscribers have a "middle of the road" pattern of performing arts consumption. They are less likely heavy book readers than non-subscribers or the general population.

Table 1 presents selected data for both extremes of non-consumption and heavy-consumption. Chi-square tests have been computed for each consumer group and each type of consumption taken by pairs. All differences are statistically significant (probability .001) except the one about book reading between the non-subscribers'

sample and the sample of general population (χ^2 test value 2.13, d.f. 5).

The sample of general population is heavily loaded with a high percentage of people who do not go out. The same sample has the highest percentage of people who read more than three books a month. Let's remember that this sample is not representative of the total population. It has been built with the same socio-demographic characteristics of the two other samples. For instance, among other traits, the level of education is similar and strongly skewed toward secondary and university education. The level of cultural interest is not the main difference between this sample and the two others. The gap between them is related to time budget, how people allocate their leisure time, are they in favor of going out or do they choose to stay home. More explanation of time budget is needed to get a better understanding of cultural consumption patterns. The two groups of subscribers and non-subscribers differ with regards to their homogeneity. The subscribers' group is very homogenous. Very few people in this group do not go to movies or do not read. At the same time, their theatre attendance does not reach the level of non-subscribers. One explanation might be that a majority of subscribers have their cultural needs fully satisfied with the theatre company performances for which they have subscribed. On the contrary among the very heterogeneous non-subscribers' sample, we have an uncommitted group with a pattern of high cultural consumption and strong desire of preserving their freedom of choice. As it can be expected we have in the same sample another group whom level of involvement with performing arts is very low. Nin percent of the non-subscribers' sample declared that their theater attendance was exceptional and due to unusual events.

TABLE 1

(% of each sample)	Subscribers	Non Subscribers	General Population
Do not go to movie-theatre	11.2	22.3	39.5
Do not go to theatre	0.6	9.4	46.3
Do not read books	5.5	10.4	11.2
Go to movie-theatre more than four times per month	7.4	10.9	5.9
Go to theatre more than ten times per year	9.5	12.4	2.9
Read more than three books per month	21.4	25.6	29.8

2. Subscribers have a greater propensity to be cultural opinion leaders than the general population; however, when their opinion leadership is compared with those of non-subscribers no significant statistical differences have been found.

Percentage differences shown in Table 2 are statistically significant between the general population and the subscribers' samples (χ^2 with probability .0001) or the non-subscribers' group (χ^2 with probability .05). It is

possible that the lack of statistical difference between subscribers and non-subscribers is due to a bias when administering the questionnaires - by mail for subscribers, by interviews within the theatre for non-subscribers. Whatsoever, the second hypothesis - H_2 - is not fully verified.

TABLE 2

(% of each sample)	Subscribers	Non Subscribers	General Population
Leaders	29.1	23.5	15.1
Non-leaders	70.9	76.5	84.9

Twenty-nine percent of the subscribers, twenty-three percent of the non-subscribers and fifteen percent of the general population viewed themselves as exercising some cultural influence on relatives and friends. Based on three distinctive cultural areas such as theatre, films and books, we look at these percentages as a measure of general opinion leadership and influence within the multi-faceted domain of culture.

TABLE 3

Own Advice Recently Asked About:	Subscribers		General Population
	Subscribers	Non Subscribers	General Population
A film	69%	56.9%	55.1%
A theatre play	60.5%	32.8%	17.6%
A book	58.2%	62.4%	60%

As it can be seen in Table 3, subscribers looked at themselves as opinion leaders about films and plays, much more than non-subscribers or the general population. Advices about books infer a similar pattern of influence when the three groups are compared.

- Cultural opinion leaders are more exposed to cultural events or more committed to cultural activities than non-leaders. This hypothesis - H_3 - is fully confirmed. Chi-square tests have been used on all differences (shown on Table 4) between leaders and non-leaders are statistically significant.

TABLE 4

% of each sample	Table 4					
	Subscribers		Non Subscribers		General Population	
	L	NL	L	NL	L	NL
Go to movie-theatres more than twice a month	39.6	30.7	57.4	35.3	41.9	21.3
Go to theatres more than five times per year.	38.3	27.5	31.9	14.4	16.2	4.6
Read more than two books per month.	55.8	38.7	68.1	43.2	81.6	43.1

- Opinion leadership overlaps exist between the three sub-categories of cultural activities taken into account in the research survey. However, this overlap varies greatly among the three groups that have been surveyed.

In Table 5 and for each three samples, the differences between observed dual leadership and estimated dual leadership have been displayed. When their differences are statistically significant, opinion leadership overlap is verified. For the subscribers' group overlap exists in all cases. Overlap is verified only in two cases with the non-subscribers' group. As for the general population sample, not only opinion leadership overlap does not exist, but in one case, movie-theatre leaders and theatre leaders, the observed dual leadership is smaller than the estimated dual leadership. Based on these three samples it seems that opinion leadership overlap is related to the cultural consumption pattern of each group. In the case where the interests of the group is strongly focused on one activity, opinion leadership overlap is unlikely. Once more it brings the question of the existence of universal opinion leaders. This study shows some evidence of the contrary. Opinion leadership overlap is verified when a group of people look at various activities as parts of a common domain. If for another group these same activities have very little in common, the overlap is not going to be verified.

- Opinion leaders look more for information in the mass media than non-opinion leaders and the influence of personal sources of information is stronger on non-opinion leaders than leaders.

Both hypotheses H_5 and H_6 are verified. This is a confirmation in the cultural and artistic domain of opinion leaders characteristics which have been well tested elsewhere.

Table 6 clearly shows the differentiation between opinion leaders and non-opinion leaders on what sources influence them most. A usual scenario would be that of an opinion leader reading in magazines the analysis of a film by one or several critics and as an intermediate expert he would transmit and influence other people looking for advice.

- Cultural opinion leaders are disseminating more cultural information than non-opinion leaders.

This hypothesis H_7 has also been verified. From previous research done by many authors on interpersonal communication and diffusion of information, we had to expect such an active behavior of cultural opinion leaders. This behavior has been approached through two questions: last time you go to the theatre, how many people have been told by you of your project? and did you try to decide people to come with you? On Table 7 all significant statistical differences between leaders and non-leaders of subscribers' and non-subscribers' samples have been reported.

- Subscribers appear to live in a closed world. Very few people like to go to the theatre alone. Going out to the theatre is a highly socialized experience. Ninety-three percent of subscribers and ninety-six percent of non-subscribers do not usually go alone to the theatre. However, these two groups of subscribers and non-subscribers rarely met. During their last theatre attendance, only six percent of all subscribers were joined by non-subscribers, and only ten percent of all non-subscribers were joined by subscribers. Finally the subscribers network of personal relations seems more

Table 5

Dual Leadership	Subscribers n = 529				Non-Subscribers n = 202				General Population n = 205			
	%	% Observed	% Estimated		%	% Observed	% Estimated		% Observed		% Estimated	
Films O.L. (1)	69	47,1 ^x	41,7		56,7	21,9	18,6		55,6	8,8	9,8	
Theatre O.L.	60,5				32,8				17,6			
Films O.L.	69	44,2 ^{xx}	40,1	24,2	56,7	41,6 ^{xx}	35,4	11,6	55,6	37,7	33,6	5,9
Book Reading O.L.	58,2				62,4				60,5			
Theatre O.L.	60,5	41,2 ^x	35,2		32,8	23,9	20,5		17,6	13,2	10,6	
Book Reading O.L.	58,2				62,4				60,5			

(1) O.L = Opinion Leaders
x Significant at p < .01 level
xx Significant at p < .05 level

Table 6

Subscribers	Media Influence	Family and Friends Influence
	%	%
Total sample	43.3	43.1
Leaders X	59.1	26.6
Non-leaders X	36.8	50.
Non-Subscribers		
Total Sample	43.9	44.1
Leaders XX	71.5	24.
Non-leaders XX	40.7	54.2

X Significant at p .001
XX Significant at p .025

Table 7

	Subscribers		Non Subscribers		
	OL (%)	NL (%)	OL (%)	NL (%)	
How many people have been told?					
Five and more	52.	63.5 (1)	40.2	60.	XX
More than five	46.1	28.	59.6	39.8	
Ask people to join you?					
Yes	72.7	54.4	61.7	45.1	XXX
No	24.	37.6	36.2	54.9	

X Significant at p < .001 level
XX " at p < .10 level
XXX " at p < .025 level

OL = Opinion Leaders
NL = Non-opinion Leaders

restricted and smaller than the one of non-subscribers. Sixty-four percent of subscribers told five persons or less about their intention to go to the theatre instead of only fifty-five percent of all non-subscribers.

Managerial Implications For The Performing Arts Organizations

Most services have to be fully experienced before consumers can really decide if their expectations are met. Of all consumer services, performing arts are the least standardized and the most influenced by situational variables. Rather than searching and processing complex information on their own, many consumers look for personal help. This explains the importance and the intensity of interpersonal communication which takes place in the cultural and artistic domain.

This situation has to be well understood by any art manager in charge of the design and the implementation of a communication policy. He has to recognize group influence which may very successfully counterbalance the influence of mass communication. Looking for the identification of opinion leaders and some way of communicating with them, is another good reason leading to the implementation of subscription policy, aside of the obvious need for loyalty and fidelity. Information from and to the subscribers is easy to get. This direct two-way of communication should be used more systematically, more creatively and more selectively. A panel of self-designated opinion leaders should be selected in order that no more than 25% of the subscribers data bank would be actively used. Not only should the season program be mailed for once, but systematic reminders of coming events should be done by mail or by phone. Creativity should be applied to both the style and the content of communication.

This communication policy with selected subscribers should be handled by art managers with a clear understanding of its limitations. One of them being the lack of openness of the subscribers' group. Opinion leaders belonging to this group are not likely to spread a cultural influence to a large number of people. It is also unlikely that they will have a natural tendency to recruit new subscribers. So the marketing plan of performing arts organizations has to recognize the existence of a de facto segmentation of subscribers and non-subscribers. The marketing mix should be adjusted to each group and specific actions should be taken for communicating with opinion leaders belonging to the non-subscribers' is very eclectic and their loyalty very difficult to obtain. However, a two sided, very informative communication should be well received by them and develop a quality image of the organization.

References

1. Bellenger D. N. and E. C. Hirschman. Identifying Opinion Leaders by Self-Report in 1977 Educators' Proceedings, Chicago, AMA, 1977, 341-44.
2. Corey, L. G. People Who Claim to be Opinion Leaders: Identifying their Characteristics by Self-Report. Journal of Marketing, 35 (October, 1971), 48-53.
3. Eiglier P. and E. Langeard. A New Approach to Service Marketing, Marketing Science Institute, Research Paper, 1977.
4. Feldman, S. and M. Spencer. The Effect of Personal Influence in the Selection of Consumer Services, in

1975 Combined Proceedings, E. M. Mazze ed. Chicago: AMA, 1975, 597-600.

5. Heitmann G. and Crocken W. E. "Theatre Audience Composition, Preferences, and Perceptions: in California Management Review: Vol. XIX, Number 2, 1976.
6. Kaali-Nagy Ch. and Garrison L. C. "Profiles of Users of the Los Angeles Music Center" in California Management Review, Vol. XV, Number 2, 1972.
7. Katz E. and P. Lazarsfeld. "Personal Influence" N.Y. Free Press, 1955.
8. Lovelock, C. H. Marketing Consumer Services: Insights from the Public and Private Sectors, in Proceedings Second Research Seminar in Marketing, Senanque, France, May 1975, Institute D'Administration Des Entreprises, ZIS, 1975.
9. Nielsen R. P. A, d MacQueen C. "Performing Arts Consumer Behavior: An Exploratory Study" in AMA Combined Proceedings, 1974.
10. Robertson, T. "Innovative Behavior and Communication" Holt, Rinehart, Winston, 1971.
11. Rogers E. M. and D. G. Cartano, "Methods of Measuring Opinion Leadership," Public Opinion Quarterly, 26, (Fall, 1962), 435-41.
12. Secretariat D'Etat A La Culture (Service Etudes et Recherches) "Les Pratiques Culturelles des Francais" 2 Tomes - Decembre 1974.
13. Tittle, C. R. and R. J. Hill, The Accuracy of Self-Reported Data and Prediction of Political Activity, Public Opinion Quarterly, 31 (Spring 1967), 103-6.
14. The Finances of the Performing Art, Vol. II: A Survey of the Characteristics and Attitudes of Audiences in Twelve U.S. Cities, Ford Foundation, 1974.

IMAGERY-ELICITING STRATEGIES:
REVIEW AND IMPLICATIONS OF RESEARCH

Kathy A. Lutz, University of Southern California
Richard J. Lutz, University of California, Los Angeles

Abstract

Recent research in cognitive psychology has focused on the use of illustration and imagery as facilitators of learning in a variety of situations. Results of this research are reviewed, with a key distinction being drawn between imagery-eliciting strategies and mental imagery. Speculations regarding the use of imagery in advertising are made in several important decision areas: advertising design and creativity, media selection and scheduling, measuring advertising effectiveness, and advertising regulation and public policy.

Introduction

The field of advertising has a long tradition of borrowing explanations for advertising effects from psychology. Among the more prominent and fruitful constructs which have been used are those associated with various learning theories, ranging from classic contiguity principles to the more currently fashionable cognitive information processing approach. A major trend in recent research in cognitive psychology has been the study of imagery and its effects on human learning and memory. This rich research tradition has remained relatively obscure to consumer researchers, although it would appear to have major implications for the study of advertising effectiveness. The purpose of the present paper, therefore, is to provide a brief review of the research on imagery in psychology, followed by a discussion of possible implications for advertising decision-making.

Review of Research on Imagery

Research on imagery has become increasingly popular recently, perhaps because the demonstrated effect of imagery is quite dramatic. Folklore suggests that a picture is worth a thousand words, yet it has been the task of empirical research to discover the full meaning and implications of this truism. In general, the literature on imagery indicates that presentations which elicit imagery are better remembered and more positively evaluated than presentations which do not. In this section, three strategies for eliciting imagery will be defined and their relationship to one another discussed. After reviewing research on each of the strategies, this section concludes with a summary and discussion of research findings.

Definition of Imagery

Imagery is defined as a mental event involving visualization of a concept or relationship. Despite its recent "re-discovery" by psychologists, mental imagery is still a rather poorly understood construct. One conception of mental imagery is physiologically based, wherein picture viewing and reported mental visualizations correspond to increased right hemisphere activity in the brain, while reading and verbal behavior correspond to a higher activation of the left hemisphere (Calloway & Harris, 1974; Galin and Ornstein, 1972; Morgan, McDonald, and MacDonald, 1971). These findings are based on studies with normal adults, but studies with brain-damaged or "split-brain" patients provide more dramatic evidence that the two sides of the brain differ. Nebes (1974) reviewed this research and con-

cluded that the two sides of the brain are not necessarily specialized for storing different types of stimuli but rather process information in different ways. The left hemisphere processes information analytically and propositionally, while the right hemisphere involves holistic or parallel processing.

The implication of brain research is that mental imagery may be better understood in terms of parallel processing than picture processing per se. The parallel processing notion is compatible with the convincing arguments and evidence that memory is propositional with one common storage format (Anderson and Bower, 1973; Norman and Rumelhart, 1976; Palmer, 1975; Pylyshin, 1973). In other words, there may be no storage distinction between visual and verbal memory, but rather the information stored as propositions or schemata could be processed in different ways. Picture processing would not, however, result in a stored representation that is identical to that resulting from processing an equivalent verbal stimulus. The differences in storage representation are caused by differences in processing rather than by different and separate storage systems. This is a difficult concept and the reader is referred to Palmer (1975) for a discussion of how mental images may be represented in a common, propositional storage system. The point is that imagery processes can affect learning and memory even though images are stored in the same type of knowledge structure that accommodates verbal information.

These speculations about the nature of mental imagery do not explain how or why mental imagery aids memory. However, the definition of mental imagery offered here emphasizes that mental imagery is a way to process information. This working definition of mental imagery is useful in interpreting and applying research results, as well as for making speculations regarding theoretical explanations of imagery processes and effects. Unfortunately, an adequate theory of imagery and visual information processing does not yet exist, despite some recent progress (Paivio, 1971; Palmer, 1975; Yuille & Catchpole, 1977).

Imagery-Eliciting Strategies

Mental imagery is generally thought to be elicited by one of the following three types of strategies, or external treatment variables: 1) pictorial stimuli such as pictures, graphs, etc.; 2) concrete verbal stimuli; and 3) imagery instructions or direct inducements. Each of these strategies will be defined and described in turn.

Most researchers accept the common dictionary definition of a picture, using their intuition to identify pictorial material. Here, pictorial material (also referred to as a picture or an illustration) is defined as any two-dimensional representation in which the stimulus array contains at least one element that is not alphabetic, numeric, or arithmetic. According to this definition, an equation or arithmetic problem would not be a picture while a word containing a "picture-letter" would be considered pictorial. One well-known example of an illustration involving a "picture-letter" is the logo for Ford Motor Company in which a light bulb replaces the 'o' in "Ford." Illustrations should also be conceptually related in some way to the topic of the communication. The effectiveness of the picture may depend on how the picture relates to the verbal information; this will be

treated later in the paper. However, pictures which are completely unrelated to the topic of the message will not be treated in this paper as they are assumed to be of no desirable consequence. For example, the use of sexy models in ads, when the model has no relationship to the product whatsoever, does not facilitate memory and may even promote negative attitudes toward the product and company (Marketing News, 1977; Steadman, 1969). On the other hand, seemingly unrelated pictures could conceivably set a mood that promotes desirable affective outcomes. For purposes of delimitation, only pictures that are conceptually relevant to the message content will be treated in this paper.

The second type of imagery strategy--concrete verbal stimuli--is perhaps the most well-defined operationally. "Imagery values" for hundreds of words have been established by having readers rate the ease with which a word arouses sensory images. These ratings are available for nouns (Paivio, Yuille, and Madigan, 1968), for verbs (Lippman, 1974), and for words commonly used by children (van der Veur, 1975). Imagery ratings are independent of other word attributes such as word frequency and associative meaningfulness, so that this variable can be studied independently of the factors that affect recall.

The two types of strategies discussed so far are examples of external stimuli that presumably affect learning and memory indirectly by eliciting internal or mental imagery. Giving instructions to the recipient to form mental pictures is a more direct way to elicit this type of mental processing. Imagery instructions refer to a statement to the learner that directs him or her to form a mental picture of the concept to be learned.

The common thread running through all three strategies or experimental treatments is the assumption that they elicit, to some degree, mental imagery, which is in turn believed to enhance memory and learning. The research findings with respect to the effects of each of the strategies on learning are now summarized.

Research on pictures. Research in this area includes studies that test the effects of

- 1) supplanting (i.e., replacing) verbal information with pictures or
- 2) supplementing (i.e., repeating) verbal information with pictures.

The findings are grouped by type of material learned in the studies including single items, paired associates, and connected discourse (i.e., sentences or stories).

Paivio (1971) reviewed studies that compared memory for pictures to memory for words. The findings indicated that pictures are remembered better than either concrete words or abstract words. Memory for visual information greatly surpasses word memory according to several more recent studies (Frdelyi and Becker, 1974; Haber, 1970; Nelson and Brooks, 1973; Paivio and Csapo, 1973). The superiority of memory for pictures has been demonstrated using advertising stimuli in an earlier study (Shepard, 1967), which showed that pictures used in magazine advertisements are easily recognized again. Comparing memory for pictures to word memory, as many of these studies did, may be inappropriate, however. The verbal equivalent of a picture is probably not a single word but rather a story or description. Since memory for words in context is better than for an unconnected word list, picture memory may not surpass verbal memory if the two types of information are equated for meaning. Another factor that may contribute to the superiority of visual memory is schematic similarity among the pictures to be remembered (Nelson, Reed and Walling, 1976). In that study, pictures that were graphically similar, were not better remembered than words. This finding suggests that other variables

should be controlled when testing the pictorial superiority effect.

Paivio (1971) also discussed studies that tested the effect of presenting a picture along with its verbal counterpart on word learning. Words were recalled better when the subject saw its accompanying picture; again, similar results have been obtained using advertising stimuli. Students recalled more company names when they had seen a pictorial depiction of the name (Lutz and Lutz, 1977). In that study, actual company names and pictures were taken from the Yellow Pages of a phone directory. Not all companies have the kind of names that lend themselves to pictorial depiction, so stimuli used in the study varied in terms of the degree of correspondence between the company name and its graphic depiction. Some pictures in the study were direct translations of the name, such as the picture for O'Bear Abrasive Saw Company which showed a bear holding a large letter "o". Other verbal-visual correspondences were not as direct. A cactus cowboy depicted Western Glass Company and Weiss Auctioneers used an owl to symbolize their name. It could be that pictures which are equivalent to their verbal counterparts facilitate memory for the company name better than pictures with a lower degree of correspondence. Perhaps future research on how verbal and visual information may be equated on the dimensions of meaning and amount of information will aid research efforts on this issue.

There are many studies on paired associated learning which indicate that a supplementary picture helps the viewer associate the two words to be learned. Learners shown a picture relating two items in a paired-associate learning task typically have significantly higher retention scores than those not presented with a picture and told only to learn the association through repetition (Davidson, 1964; Kerst and Levin, 1973; Lippman and Shanahan, 1973; Reese, 1965). One necessary condition for the facilitative effect is the interactive nature of the picture. An interactive picture integrates the two items in some mutual or reciprocal action. When the picture is not interactive and the items are depicted or imagined side by side, paired associate learning is not necessarily facilitated (Bernbach and Stalonas, 1973; Bower, 1972; Neisser and Kerr, 1973).

Interactive pictures facilitate memory for ads according to a study mentioned earlier, which used company-product pairs and their accompanying pictures from ads in the Yellow Pages of a telephone directory (Lutz and Lutz, 1977). The group receiving the interactive imagery treatment remembered significantly more company names than the non-interactive imagery group. Also, while the interactive imagery group remembered more brand names than its control group who saw only the company-product names with no accompanying pictures, the non-interactive imagery group did not differ significantly from its control group.

The interactive imagery group in the Lutz and Lutz (1977) study received two types of pictures, including picture-interaction items and letter accentuation pictures. In the latter type, some letter or letters in the brand name are made to look like the product or one of its characteristics. The logo for Mullin Lumber Company, in which the letters of "Mullin" are drawn as wood grain boards, is an example of letter accentuation. Picture-interaction items unite a depiction of the brand and of the product or service in one picture. An example is the logo for Jack's Camera Shop which shows a playing card jack holding a movie camera to his eye. Students remembered picture interaction items better than letter accentuation pictures, which was also the finding in an earlier study (Lippman and Shanahan, 1973).

One dimension of an interactive picture that has not yet been investigated is the "dynamism" or animation of the

interaction. Interactive pictures may be either dynamic or static, with dynamic interactions probably being the more memorable type. Television ads provide examples of dynamic interactive pictures in which the interaction is animated. Research on attention and reinforcement indicates that a visual stimulus change in the environment elicits attention to that action by the viewer (Kish, 1966). So the viewer's attention can be drawn to the important information by animating the interaction.

Presenting pictures along with stories and other forms of connected discourse is also facilitative. As mentioned before, pictures can accompany verbal material by supplanting or replacing part of it and by supplementing the verbal information. Although findings are not conclusive, some studies indicate that learning and attitude are positively enhanced when pictures replace some of the verbal information (Booher, 1975; Rigney and Lutz, 1976). When pictures completely replace the verbal text, however, comprehension may not be facilitated. Children who received a story in the form of a series of pictures remembered no more than students who read the printed text (Levin, 1973). Levin speculated that the pictorial treatment would have been more beneficial with some accompanying verbal labels. Another explanation of Levin's result is that his verbal comprehension test may have failed to test the kind of knowledge acquired by the pictorial treatment.

Most investigations in the literature on illustration have tested the effect of supplementing verbal material with pictures and have assessed outcomes with verbal posttests. In an earlier review, Samuels (1970) concluded that "pictures, when used as adjuncts to the printed text, do not facilitate comprehension." However, several recent studies indicate that supplemental pictures facilitate learning under some circumstances at a variety of age levels. Recent findings indicate that children learn more from a prose passage when it is accompanied by pictures (Bender and Levin, 1976; Lesgold, Levin, Shimron and Guttman, 1975; Rohwer and Harris, 1975; Rohwer and Matz, 1975). Illustrations also aid children's memory for more difficult material that contains new concepts (De Rose, 1976). Seeing pictures apparently helps adults and high school students learn new concepts (Dwyer, 1972; Holliday, 1975). In Dwyer's study, pictures aided performance on graphic or drawing tests but not on verbal learning measures, indicating that pictures may result in non-verbal outcomes.

One recent study, however, found no facilitative effect for supplementary illustrations with adult populations. College students who read a passage about revolutions did not benefit from supplemental drawings, but mental imagery instructions did facilitate their performance (Rasco, Tennyson, and Boutwell, 1975). A second study yielded a similar outcome with high school seniors. Perhaps, when the visual information and verbal information are redundant and the verbal information is easily understood, illustration may not be beneficial for adult learners.

Research on pictures also indicates that learning is improved when students are asked to draw their own illustrations. Subject-generated pictures facilitated paired-associate and word learning (Bull and Wittrock, 1973; Wittrock and Goldberg, 1975), story comprehension (Lesgold, Levin, Shimron, and Guttman, 1975, Snowman and Cunningham, 1975), and recall of new concepts (Dansereau et al., 1975; Rasco, Tennyson, and Boutwell, 1975).

Although many ads are presented in a story format with accompanying pictures, no studies have empirically tested the differential effects of supplementing or supplanting the story line with pictures.

Research on concrete verbal stimuli. Research findings on the imagery value of verbal stimuli generally indicate that concreteness facilitates memory for words, paired-associates, and connected discourse. Studies will again be grouped around these three stimulus categories.

Paivio (1969) reviewed the effects of concreteness or imagery value on word recall. He cited a series of experiments that varied imagery value and word meaningfulness independently (or partialled out meaningfulness, and found that word recall was positively related to the level of word concreteness. More recent studies have confirmed the finding that words high in imagery value are remembered better than words of low imagery value when other factors affecting recall such as meaningfulness and frequency are held constant (Craig, 1973; Elliott, 1973; Griffith and Johnson, 1973; Paivio and Csapo, 1973; Wortman and Sparling, 1974).

No published studies have yet investigated the effects of word concreteness on recall of brand or company names. According to the research evidence, names that are concrete and picturable should be more memorable than abstract words, proper names, or abbreviated names.

Concreteness also affects paired-associate learning. When words are to be associated, memory for the pair can be affected by concrete connective words. Several studies varied the kind of connective words used to relate word pairs and found that verbs were more effective in facilitating memory for the pair than conjunctions or prepositions (Reese, 1965; Rohwer, Lynch, Levin and Suzuki, 1967). Although imagery values for the connecting words were not compared, it seems quite likely that the verbs had much higher imagery values than the conjunctions and prepositions used in those experiments. Other results suggest that embedding the word pair in the context of a sentence aids memory for the pair (Davidson, 1964; Levin, Davidson, Wolff and Citron, 1973; Lippman and Shanahan, 1973; Milgram, 1967). The sentences related the word pairs with action verbs, which probably made the association of the pair more concrete and picturable.

The use of concrete connectives in advertising messages has not yet been investigated. The research evidence suggests that messages which use concrete words to connect company name and product claims should make the messages more memorable. Assuming that action verbs are more concrete than linking verbs (although this has yet to be established empirically), the more effective ad will tell what the product does (its usage benefits) rather than what the product is (its attributes).

Connected discourse such as sentences prose passages, and stories, is better remembered if it contains concrete words rather than words of low imagery value (Anderson, 1974; Johnson, Bransford, Wyberg, and Cleary, 1972; Montague & Carter, 1973; Yuille and Paivio, 1969). However, words used in advertising stories and vignettes have not been systematically varied to test the thesis that concrete words would make the brand name and product claims more memorable.

Research on imagery instructions. The results of numerous studies on mental imagery instructions are reviewed and discussed by Paivio (1971) and will be only briefly summarized here. Giving learners imagery instructions facilitates paired-associate and word learning and the effects are much like those reported in the previous section on pictures. The interactive feature of the resultant image remains a necessary condition for maximum effectiveness. Learners told to mentally imagine two objects in interaction remember more word pairs than learners told to mentally picture the objects side by side or on opposite sides of a room.

One advantage of interactive imagery, as opposed to experimenter-supplied interaction pictures, is that the interaction can be idiosyncratic and bizarre, which has been recommended by ancient as well as modern day mnemonists (Lorayne & Lucas, 1974). The claim that bizarre images facilitate memory was not supported, however, by a recent empirical study (Nappe & Wollen, 1973). In fact, students in that study took longer to form bizarre mental images than "common" images. Although bizarreness does not appear to be facilitative, the related attribute of uniqueness or distinctiveness may be an effective attribute of facilitative mental images. A recent study varied the level of schematic and conceptual similarity of pictures and found that visual recall was no better than word recall under conditions of high schematic similarity (Nelson, et al., 1976). These results suggested that pictures or mental images which are too common may be confused with other visual information and may not be remembered.

Although there are no studies on the effects of imagery instructions on associating brand names with product claims in advertising, one particular ad provides an example of how to induce a common yet distinctive image. A Johnston's Yogurt ad directs the listener to "look for the sunny yellow cups" which probably elicits a mental image that is easy to form and yet is distinctive from other, similar products (unless other yogurt manufacturers use yellow containers as well).

Recent findings indicate that imagery instructions also facilitate memory for connected discourse. Children remember more after reading a passage when they receive imagery instructions, especially if the instructions are preceded by training in the use of this strategy (Kulhavy and Swenson, 1975; Lesgold, McCormick and Golinkoff, 1975; Pressley, 1976). Children may not benefit from imagery instructions, however, if the material to be learned is new or difficult (De Rose, 1976).

The pattern of effects is similar for adults. Imagery instructions facilitate memory for easy material such as sentences and narratives (Anderson, 1971; Anderson and Hiddi, 1971; Rasco et al., 1975). Imagery instructions per se appeared to have no effect on memory in one study, but closer analysis reveals that mental imagery did facilitate learning (Anderson and Kulhavy, 1972). Students in that study who reported that they learned the text passage by using "mental pictures" remembered more than students who reported not using mental imagery. The findings of this study are important because, while the nominal treatment of imagery instructions may fail to elicit mental imagery, reported use of mental imagery does correspond to increased recall. Another study showed no overall benefit from imagery instructions during a limited time period (Lesgold, Curtis, De Good, Golinkoff, McCormick and Shimron, 1974). However, adults who received the imagery instructions remembered more from the first half of the passage than did control subjects, suggesting that the imagery strategy required more time but was beneficial. The authors noted that imagery should not be considered a useful strategy if it produces no learning gains per unit time.

Directing adults to use mental imagery while learning new concepts may not facilitate learning, according to recent findings (Rigney & Lutz, 1976). The failure of mental imagery instructions to aid new concept learning in both adults and children is plausible because it would not be expected that people are capable of forming mental images of concepts with which they are not yet familiar.

One problem with trying to induce the learner to use mental imagery while reading is that the act of reading tends to interfere with imagery production, because both activities presumably involve some visual information processing. Findings by Levin and Divine-Hawkins (1974)

indicated that imagery instructions were more facilitative when children listened to a passage rather than read it. Brooks (1967) also found that mental imagery was more readily elicited during listening than reading by adults. Most of the studies on imagery instructions told the subject to visualize while reading, so the results of those studies may have been even stronger if the subjects had listened to, rather than read, the verbal information.

Telling an advertising audience to form mental imagery while hearing or reading an advertising message has not been empirically tested, but the literature on this treatment suggests that it could facilitate memory for the ad. Imagery instructions may be especially effective with auditory media such as TV and radio, which would allow listeners to form idiosyncratic mental images while listening to the message.

Discussion and Summary

Outcome measures. The effects of imagery strategies must be viewed relative to the way learning outcomes are assessed. Tests of comprehension are most often only verbal, meaning that only verbal outcomes of presentation mode are assessed. Several studies tested the effects of pictorial treatments with graphic retention measures, which tapped knowledge not assessed by the verbal posttests (Dwyer, 1972; Lutz and Rigney, 1977).

Recent evidence on left and right hemispheric processes of the brain suggests that the left brain often dominates in processing and retrieving verbal or propositional information while the right hemisphere processes and retrieves information holistically. Levy, Trevarthen, and Sperry (1972) presented a dramatic demonstration of this difference when they showed split-brain patients a chimerical picture such as a face of which each half is actually from a different face. Using tachistoscopic methods, each half of the picture is sent to a different hemisphere and because of the nature of the patients' surgery, the hemispheres cannot communicate as in the normal brain. Results indicate that when the patient is asked to point to the face that he saw (i.e., recognize it), he points to the face corresponding to the half sent to the right hemisphere with no realization that the presented stimulus was composed of two different faces. When the patient is asked to verbally describe the features of the face that was shown (i.e., recall it), however, the face corresponding to the half projected to the left hemisphere is retrieved and described. This is convincing evidence that the left hemisphere is analytical and propositional while the right is holistic. But, more important to the point of assessing outcomes of treatments, this study demonstrates that the retrieval task may determine the nature of the information retrieved. Although this evidence comes from patients with altered brain function, normal individuals may display recall and recognition differences to some degree.

In assessing advertising effects, asking consumers to recall and state ad claims is probably a left hemisphere task. In contrast, a recognition test requiring the consumer to identify illustrations used in an ad may elicit more right hemisphere processing. The nature or type of information retrieved using these two approaches may differ. One recent study suggests that graphic assessment techniques reveal information stored in memory that is not necessarily retrieved by verbal measures (Rossiter, 1976).

Most of the outcomes assessed by the studies reviewed in this section dealt primarily with cognitive rather than affective outcomes of imagery-eliciting strategies. A few of the studies assessed the effects of imagery on attitude and found that imagery-eliciting strategies seem to promote positive attitudes in addition to

facilitating learning. Samuels (1970) concluded in his earlier review on imagery effects that although pictures may not always affect cognitive outcomes, they do seem to facilitate favorable affective outcomes. More research is needed to clearly establish the effects of imagery on attitude.

Imagery-eliciting strategies may also affect decision making processes. Some evidence suggests that people weight concrete examples more heavily than abstract information when making judgments or decisions (Nisbett, et al., 1976). Thus outcomes other than learning alone should be considered when determining the effects of imagery-eliciting strategies.

Summary of facilitative conditions. Although findings are mixed, the literature on imagery generally indicates that pictures, concrete words, and mental imagery instructions facilitate memory and may enhance attitude as well. Individual differences in prior knowledge may influence the size of the imagery effect. Pictures may not help a reader who already knows a lot about a topic. For example, adults in the Rasco et al. study (1975) did not benefit from seeing pictures. These readers did benefit, however, by forming their own mental images while reading. The opposite pattern of results was found in a study with children who were learning material intended to be difficult or instructional. The children remembered more when adjunct pictures accompanied the text than when they were given mental imagery instructions (De Rose, 1976).

These seemingly conflicting results may be explained by an interaction hypothesis that mental imagery is facilitative under conditions of prior familiarity while pictures are more beneficial when the learner has less prior knowledge. There are several plausible arguments favoring this hypothesis. People should not be capable of forming mental images of concepts that they have not yet acquired, so a mental imagery strategy would be rather ineffective for new concept learning. Pictures, on the other hand, do aid learning of difficult material but may not make any difference when an adult reads prose text that is easy to comprehend.

Another way to look at the differential effectiveness of these strategies is that learners may benefit more from their own strategies than imposed strategies, under certain conditions. Bobrow and Bower (1969) demonstrated that subject-generated verbal mediators were superior to experimenter-provided verbal mediators in a paired-associate learning task. Since that study, other researchers have turned attention to the potential value of instructing or training learners to apply their own learning strategies. Picture drawing and mental imagery generation are examples of learner-generated strategies that appear to facilitate learning when the learner is capable of performing these tasks.

Other variables that may affect the strength of the imagery effect include characteristics of the picture such as figural interaction and dynamism and the relationship between the verbal and visual information in a message. Pictures must be related to but not always repetitive of passage content to be facilitative. Adults who are good readers may not benefit from viewing a supplementary picture when they read simple narratives or familiar material. When the illustration presents additional information, however, picture viewing results in more effective and efficient learning. Children or learners with little prior knowledge seem to benefit from illustrations whether they supplement or supplant some of the verbal information. Hence, the learner's prior knowledge may interact with the relationship of verbal and visual information.

A consideration of the principles and effects of imagery is relevant to a number of important decision variables in the advertising mix. While little empirical research on imagery and advertising has been undertaken to date, the perspective provided by the imagery literature can, nevertheless, be useful in structuring advertising decision-making along lines consistent with imagery principles.

The following discussion of advertising implications is not intended to be exhaustive and conclusive, but rather speculative and limited to the areas showing the most obvious potential benefits of adopting an imagery perspective. The areas to be considered include the use of imagery in 1) advertising design and creativity, 2) media selection and scheduling, 3) measuring advertising effectiveness, and 4) advertising regulation and public policy.

Advertising Design and Creativity

Imagery has its most obvious implications in the area of advertising design. For instance, the use of pictures in ads is hardly a new idea to advertisers, who have relied on illustrations for years in the quest for more interesting and persuasive advertising campaigns. However, the principles of imagery suggest a richer perspective on the problem of using pictures to communicate.

Some illustrations are more effective than others. For instance, the use of interactive imagery, wherein a brand name is pictorially portrayed in relation to a product concept, should result in increased association of the brand and its attributes or benefits as demonstrated by Lutz and Lutz (1977). In contrast, a non-interactive image portraying either the brand or its benefits (or perhaps both, but not in a single, integrative picture) may result in retention of the pictorially presented stimulus, but not the critical relationship between the brand and its benefits. Krugman (1978) has argued for the use of "memorable" illustrations as an aid to effective advertising. The concept of interactive imagery suggests that it is not enough that the illustration be memorable but that it must also depict the relationship between the brand and its proffered benefits. A Los Angeles exterminator named Arrow Pest Control uses as its logo an insect pierced by an arrow. Remembering the illustration also means remembering the brand name (Arrow) and its product offering (bug killing). A cursory examination of much contemporary advertising suggests that many advertisers conjure up dramatic illustrations which do little to communicate the brand-benefit linkage.

From an imagery perspective, then, appropriate illustrations can vividly portray information - i.e., the concepts that the advertiser wishes to communicate. When pictures do not communicate relevant information, they should not be expected to enhance advertising effectiveness, except perhaps indirectly, as under the distraction hypothesis (Bither, 1972; Wright, 1973). One important issue in treating illustrations as information presentations is whether they should serve as supplements to, or as substitutions, for, verbal presentations of the same information. While the former use would imply some redundancy in the commercial message, this redundancy may constitute a type of repetition within the ad, which is presumably a desirable feature. Additionally, the use of pictures in lieu of verbal copy claims may be hazardous, in that audience interpretations of pictures may be more varied than comprehension of verbal claims. A picture may indeed be worth 1000 words, but it may not be the same 1000 words for all members of the audience. The use of concurrent verbal stimuli may help to homogenize response to the illustration while preserving the increased vividness gained by use of the picture.

Thus far, the use of imagery in advertising design has focused on external imagery in the form of pictures and illustrations. Pictures are assumed to be facilitative because they elicit internal or mental imagery. Mental imagery may be induced through means other than the presentation of pictures, however. The use of concrete terms, verbal descriptions of pictures, and imagery instructions are other ways to elicit mental imagery in the mind of the potential consumer.

An interesting consideration in the use of nonpictorial spurs to mental imagery is the possibility that there are two operators at work in the process. One operator is the ability of the consumer to come up with an appropriate image without being shown one by the advertiser. It would seem likely that this ability is not equally present in all audience members and that some will be "lost" at this step. The second operator conceivably at work is the idiosyncratic nature of the images that are created. The power of the human imagination is vast and may supercede any advertiser-provided stimulus in being personally relevant or vivid to the consumer. Thus, for those consumers who do form mental images, the resultant images may be much more powerful than those stimulated by an illustration.

One may question whether consumers are sufficiently involved with commercial advertising to actually devote time and effort to forming a mental image. A similar criticism applies to the cognitive response model of communications effects (Wright, 1973) to which an ingenious, if risky, counterploy by an advertiser might be to instruct the consumer to "think about it" (i.e., generate cognitive responses) (Ray, 1977). A similar solution to the mental imagery problem may involve telling the consumer to "imagine yourself" in a particular situation or performing a certain behavior.

Media Selection and Scheduling

Viewing the communications task from an imagery perspective immediately suggests that certain media vehicles are more suited to certain kinds of imagery uses. The various media have different strengths and weaknesses which can be taken into consideration in planning a campaign relying on imagery principles. One basic distinction which has been made is that between the broadcast media (TV, radio) and print media (magazines, newspapers). The former are presumably more "intrusive" but characterized by less audience involvement than the latter (Krugman 1965), a source of concern to many advertisers. Consideration of imagery principles may help advertisers to overcome or perhaps even benefit from low involvement on the part of consumers.

Television. This medium is especially suited to the use of dynamic interactive imagery, thus making full use of the power of illustrations in affecting learning, attitudes, and behavior. The use of action in the commercial can focus the viewer's attention on the critical interaction between the brand and the product or service and make it much more memorable. An example of this type of imagery is the ad for Hunt's tomato sauce in which thick tomato sauce is shown being poured into the "u" in Hunt's. This is an example of a letter accentuation image, which is useful for brand names that cannot be depicted pictorially. An example of dynamic imagery that may not direct attention to the brand-benefit association is the ad for AMC that shows a number of interesting pictures appearing next to a moving, flashing line. The line holds the viewer's attention very well, but the viewer may fail to notice what brand was advertised. Thus, the ad may not direct the viewer's attention to the critical information.

Another imagery technique made possible by the capability of TV is modeling. Viewers can be shown the actual use of a product, thus portraying the behavior desired of

the audience, as well as communicating beliefs and cognitions about the product.

As noted previously, viewers are often described as having "low involvement" when viewing TV ads. Viewing pictures as opposed to words seems to require much less involvement or cognitive processing for perception and comprehension. As suggested by Krugman (1978), right hemispheric processing may be at work. This can work to the advertiser's advantage, since it takes less effort to attend to and comprehend a picture than is required to understand an equivalent amount of information presented verbally. Thus the use of pictures and models to create interactive, highly memorable images is especially important in TV advertising since this is generally considered to be a low involvement medium.

Radio. While the use of imagery in radio advertising has certain limitations inherent to the medium (i.e., low involvement and inability to present information pictorial), creative inducements to use mental imagery could make radio ads as effective if not more effective than media which present external imagery. The literature on mental imagery indicates that subject-generated imagery may be superior to picture viewing for adult learners. Lorayne and Lucas (1974) in their popular The Memory Book, claim that bizarre, unique, individualized mental images are the most effective. Instructions to use individual mental imagery may result in not only improved memory for the message, but also increased involvement by the listener.

There are several ways to induce imagery on radio. One is to use concrete language to describe a picture, hopefully inducing the listener to form a mental image. The description should ideally be of an interactive image. Perhaps the most effective approach would be to coordinate radio descriptions with dynamic interactive pictures that are presented in TV ads. A description of the image may elicit memory for the already viewed dynamic image so that the listener "sees" the picture again in the "mind's eye".

Another way to elicit imagery is to give imagery instructions to the listener either directly or implicitly, telling the listener to picture themselves in a situation. An example of imagery instructions in a radio commercial is the Delta Airlines ad that directs listeners to "raise your hand, look at your watch, and repeat after me: Delta is ready when you are." The listeners mentally picture themselves looking at their watches and thinking about Delta Airlines' attribute of being on time.

Print. Magazines and newspapers are assumed to enjoy higher involvement on the part of their readers so ads in these media can safely devote more of the communications task to verbal information. The relationship between the verbal and visual information presented in print media may be different than in media with lower audience involvement. Pictures can supplement the main points made verbally.

Also a particular advantage in the print media as a result of higher reader involvement is the possibility of providing mental imagery instructions with somewhat more assurance that some portion of the audience will actually attempt to form mental images. Another related possibility would be to leave space in the ad for readers to draw a picture of themselves using the product. A modified picture-generation task such as "connect the dots" may be successful. Even the covert response of thinking about drawing the picture may help to crystallize an image in the consumer's mind.

Outdoor. Outdoor advertising or billboards are mentioned here only briefly as a medium which can benefit greatly from the use of imagery principles. Severely restricted in space and exposure time, a billboard will be most

effective by arousing some sort of mental imagery in the consumer, either through the use of illustrations or brief imagery instructions. It is also noted in passing that certain types of mechanical billboards may take advantage of dynamic interactive imagery, although this would be the exception rather than the rule.

Media mix. In selecting the media mix for a complete advertising campaign, integrated use of imagery-eliciting strategies in the various media should be one of the objectives. For instance, the more involving print media might use concrete language and imagery instructions, with the broadcast media serving the role of eliciting a reconstruction of those images. As another example, television advertising might utilize dynamic interactive pictures, with static portrayals of the same picture appearing in print form. The basic notion is to use each media vehicle to its own particular advantage, perhaps shifting the mix to match the type of imagery appeal being used. A campaign emphasizing dynamic interactive imagery may use more TV initially, followed by a shift into more print media and radio spots.

Media scheduling. One of the primary advantages offered by the use of imagery-eliciting strategies in advertising is the improved efficiency of communication. Presumably the same advertising effects can be achieved with fewer exposures to imagery-eliciting ads than their non-image counterparts. Germane to this point, Tversky and Kahneman (1973) discuss the so-called availability hypothesis: "any incident that makes the occurrence of an event easy to imagine or recall is likely to enhance its perceived frequency". This hypothesis has obvious implications for the problem of ad wearout. If an image-eliciting ad is easier to recall, consumers are likely to believe that they have seen the ad more often. Thus, the ad may have a shorter "life cycle" before becoming an annoyance to consumers, unless the ad is scheduled less frequently to counteract this effect. Less frequent scheduling of the ad means improved cost effectiveness.

Measuring Advertising Effectiveness

Hierarchy-of-effects. Perhaps one of the most controversial issues in advertising evaluation revolves around the notion of the "hierarchy-of-effects" (e.g., Lavidge and Steiner, 1961). Critics have assailed the hierarchy as inaccurate in its portrayal of response to low involvement advertising (e.g., Krugman, 1965), and others have suggested that as many as three different forms of hierarchy may exist (Ray, 1977). Imagery may be an important determinant of the power of a hierarchy-of-effects approach to modeling communications effects.

In its simplest form, the hierarchy consists of three steps: 1) cognition, or beliefs about the brand, 2) affect, which represents attitude toward, or evaluation of, the brand, and 3) conation, representing behavior with respect to the brand. For "high involvement" products, the steps are believed to occur in the order listed above, while in the "low involvement" case, Steps 2 and 3 are thought to be reversed (Ray, 1977). If this reversal is in effect, then traditional forms of advertising pretesting which rely on brand attitude as a surrogate for subsequent market behavior may be quite useless. Consumers are viewed as being in a very passive role, learning "without involvement" (Krugman, 1965) and not forming any true evaluation of the brand until after initial purchase.

The use of effective imagery strategies in advertising may act to 1) lift the low involvement situation to a higher level of involvement, 2) promote non-verbal cognitions and 3) elicit attitudes toward that brand, thus making the cognition-affect-conation hierarchy a viable one for assessing advertising effects. Pictorial treatments are associated with more positive attitude

toward the message, according to several studies reviewed earlier. Nisbett, et al. (1976) cite several studies showing that concrete information is much more influential in human judgmental processes than is more abstract information. It is not unreasonable to expect that the very presentation of concrete information in an ad may stimulate the sorts of judgmental processes upon which the hierarchy-of-effects model rests.

Recall vs. recognition. A related issue in advertising evaluation is the appropriate measure of consumer learning outcomes. Krugman (1965, 1978) has argued for the use of recognition rather than the more traditional verbal recall. The importance of using visual in addition to verbal assessment techniques after picture viewing has been recognized recently in consumer research. Rossiter (1976) reasoned that product information perceived and retrieved visually may differ considerably from attribute information retrieved verbally. Children at several age levels in his study appeared to have much information about cereals available when drawing a picture of the product that did not coincide with verbal recall. More importantly, this visual information also influenced their preferences to some extent.

The use of imagery in advertising may engage the brain more fully in the response to the ad. Both the left and right hemispheres probably become more activated in processing a combination of verbal and pictorial information. Ad testing technology must keep pace by using procedures adequate for assessing the effects of the image-related portion of the ad.

Advertising Regulation and Public Policy

Recent consumer-oriented public policy has heavily emphasized the need for advertising to serve an informational function for the consumer. The essential thrusts of affirmative disclosure, nutrient labeling, and advertising substantiation all rest on the notion of the consumer as an information processor. Unfortunately, most of the focus has tended to be on verbally presented information, rather than on visually presented information, despite the fact that the FTC is concerned about "implied" product claims.

As noted by Eighmy (1978), much of the work being done presently in public policy relies on the use of copy testing, which has as its main focus verbal copy claims, and to a lesser extent the implications of a visual context. Yet good imagery in advertising may be much more powerful than any verbal claim, in part due to the former's concreteness (Nisbett, et al., 1976). Thus, public policy directed at advertising should immediately shift greater attention to the imagery in advertising in order to more adequately perform its job of protecting/informing consumers. The prevailing view of the consumer as an information processor must be broadened to encompass not only left hemispheric verbal processing, but also the more subtle yet perhaps more powerful right hemispheric processing of visual imagery.

Conclusion

This paper has attempted to summarize briefly the research on illustration and imagery which has appeared primarily in the psychology literature. Implications of these findings for research on, and the practice of, advertising would seem to be rather far-ranging. A number of potential issues have been raised for the application of imagery principles to advertising decisions. Further research is needed to determine which areas of advertising can benefit most from an imagery perspective, and how communications in general can be improved through the use of imagery.

References

- John R. Anderson and Gordon H. Bower, Human Associative Memory (Washington D.C.: Winston, 1973).
- Richard C. Anderson, "Encoding Process in the Storage and Retrieval of Sentences," Journal of Experimental Psychology, 91 (December, 1971), 338-340.
- Richard C. Anderson, "Concretization and Sentence Learning," Journal of Educational Psychology, 66 (April, 1974), 179-183.
- Richard C. Anderson and Janet L. Hidde, "Imagery and Sentence Learning," Journal of Educational Psychology, 62 (December, 1971), 526-530.
- Thomas C. Arnold and Francis M. Dwyer, "Realism in Visualized Instruction," Perceptual and Motor Skills, 40 (1975), 369-370.
- Bruce G. Bender and Joel R. Levin, "Why Do Pictures Facilitate Young Children's Learning of Oral Prose?" Paper presented at the meeting of the American Educational Research Association Annual Meeting, San Francisco, April, 1976.
- Harley A. Bernback and Peter M. Stalonas, "The Effect of Isolated Imagery on Free Recall," Bulletin of Psychonomic Science, 1 (1973), 227-228.
- Stewart W. Bither, "Effects of Distraction and Commitment on the Persuasiveness of Television Advertising," Journal of Marketing Research, 9 (February, 1972), 1-5.
- Samuel A. Bobrow and Gordon H. Bower, "Comprehension and Recall of Sentences," Journal of Experimental Psychology, 80 (June, 1969), 455-461.
- Harold R. Booher, "Relative Comprehensibility of Pictorial Information and Printed Words in Proceduralized Instruction," Human Factors, 17 (June, 1975), 266-277.
- Gordon H. Bower, "Imagery as a Relational Organizer in Associative Learning," Journal of Verbal Learning and Verbal Behavior, 9 (1970), 529-533.
- Gordon H. Bower, "Mental Imagery and Associative Learning," In L. Gregg (Ed.) Cognition in Learning and Memory (New York: Wiley and Sons, 1972).
- L.R. Brooks, "The Suppression of Visualization in Reading," Quarterly Journal of Experimental Psychology, 19 (1967), 289-299.
- Britta L. Bull and Merlin C. Wittrock, "Imagery in the Learning of Verbal Definitions," British Journal of Educational Psychology, 43 (1973), 289-293.
- E. Calloway and P.R. Harris, "Coupling Between Cortical Potentials from Different Areas," Science, 183 (1974), 873-875.
- Ellis M. Craig, "Role of Mental Imagery in Free Recall of Deaf, Blind, and Normal Subjects," Journal of Experimental Psychology, 97 (February, 1973), 249-253.
- Donald F. Dansereau et al., Effective Learning Strategy Training Program: Development and Assessment. AFHRL-TR-75-41. Lowry AFB, Colo.: Technical Training Division, Air Force Human Resources Laboratory, June, 1975.
- Robert E. Davidson, "Mediation and Ability in Paired-Associate Learning," Journal of Educational Psychology, 55 (1964), 352-356.
- Thomas M. DeRose, "Pictorial and Verbal Strategies in Children's Reading Comprehension." Paper presented at the meeting of the American Education Research Association, San Francisco, April 1976
- Francis M. Dwyer, "The Effect of Overt Responses in Improving Visually Programmed Science Instruction," Journal of Research in Science Teaching, 9 (January, 1972), 47-55.
- John Eighmey, "A Perspective on Advertising Copy Research and Advertising Law Enforcement," in J.C. Maloney and B. Silverman (eds.), Attitude Research Plays for High Stakes (Chicago: American Marketing Association, 1978), forthcoming.
- Lee Elliott, "Imagery Versus Repetition Encoding in Short-term and Long-term Memory," Journal of Experimental Psychology, 100 (October, 1973), 270-276.
- Matthew H. Erdelyi and Joan Becker, "Hypermnesia for Pictures: Incremental Memory for Pictures but not Words in Multiple recall trials," Cognitive Psychology, 6 (1974), 159-171.
- David Galin and Robert Ornstein, "Lateral Specialization of Cognitive Mode: An EEG Study," Psychophysiology, 9 (1972), 412-418.
- Douglas Griffith and William A. Johnson, "An Information-Processing Analysis of Visual Imagery," Journal of Experimental Psychology, 100 (September, 1973), 141-146.
- Ralph N. Haber, "How We Remember What We See," Scientific American, 222 (1970), 104-112.
- William G. Holliday, "The Effects of Verbal and Adjunct Pictorial-Verbal Information in Science Instruction," Journal of Research in Science Teaching, 12 (1975), 77-83.
- Marcia K. Johnson, John D. Bransford, Stanley E. Nyberg, and John J. Cleavry, "Comprehension Factors in Interpret-and Memory for Abstract and Concrete Sentences," Journal of Verbal Learning and Verbal Behavior, 11 (1972), 451-454.
- Stephen Kerst and Joel R. Levin, "A Comparison of Experimenter-provided and Subject-generated Strategies in Children's Paired-associate Learning," Journal of Educational Psychology, 65 (December, 1973), 300-303.
- G.B. Kish, "Studies of Sensory Reinforcement," In W.K. Honig (ed.), Operant Behavior: Areas of Research and Application (New York: Appleton-Century Crofts, 1966), 109-159.
- Herbert E. Krugman, "The Impact of Television Advertising: Learning Without Involvement," Public Opinion Quarterly, 30 (1965), 583-96.
- Herbert E. Krugman, "Low Involvement Theory in the Light of New Brain Research," in J.C. Maloney and B. Silverman (eds.), Attitude Research Plays for High Stakes (Chicago: American Marketing Association, 1978), forthcoming
- R.W. Kulhavy and I. Swenson, "Imagery Instructions and the Comprehension of Text," British Journal of Educational Psychology, 45 (1975), 47-51.
- Robert I. Lavidge and Gary A. Steiner, "A Model for Predictive Measurements of Advertising Effectiveness," Journal of Marketing, 25 (1961), 59-62.
- Allan M. Lesgold, M.E. Curtis, H. DeGood, R.M. Golinkoff, C. McCormick, and J. Shimron. The Role of Mental Imagery in Text Comprehension: Preliminary Studies. (Pittsburgh: University of Pittsburgh, Learning Research and Development Center, 1974).

- Alan M. Lesgold, Joel R. Levin, Joseph Shimron, and Joseph Guttman, "Pictures and Young Children's Learning from Oral Prose," Journal of Educational Psychology, 67 (October, 1975), 636-642.
- Alan M. Lesgold, Claire McCormick, and Roberta M. Golinkoff, "Imagery Training and Children's Prose Learning," Journal of Educational Psychology, 67 (October, 1975), 663-667.
- Joel R. Levin, "Inducing Comprehension in Poor Readers: A Test of a Recent Mode," Journal of Educational Psychology, 65 (August, 1973), 19-240.
- Joel R. Levin, Robert E. Davidson, Peter Wolff, and Michael Citron, "A Comparison of Induced Imagery and Sentence Strategies in Children's Paired-associated Learning," Journal of Educational Psychology, 64 (June, 1973), 306-309.
- Joel R. Levin and Patricia Divine-Hawkins, "Visual Imagery as a Prose-learning Process," Journal of Reading Behavior, (1974), 23-30.
- Joel R. Levin, Elizabeth D. Ghatala, Thomas M. DeRosa, Larry Wilder, and Richard W. Worton, "A Further Comparison of Imagery and Vocalization Strategies in Children's Discrimination Learning," Journal of Educational Psychology, 67 (February, 1975), 141-145.
- Jerre Levy, Colwyn Trevarthen, and R.W. Sperry, "Perception of Bilateral Chimeric Figures Following Hemispheric Deconnection," Brain, 95 (1972), 61-78.
- Marcia Z. Lippman, "Enactive Imagery in Paired-associate Learning," Memory & Cognition, 2 (1974), 385-390.
- Marcia Z. Lippman and Morris W. Shanahan, "Pictorial Facilitation of Paired-associate Learning: Implications for Vocabulary Training," Journal of Educational Psychology, 64 (April, 1973), 216-222.
- Harry Lorayne and Jerry Lucas, The Memory Book (New York: Stein and Day Publishers, (1974).
- Kathy A. Lutz and Richard J. Lutz, "The Effects of Interactive Imagery on Learning: Application to Advertising," Journal of Applied Psychology, 62 (August, 1977) 493-498.
- Kathy A. Lutz and Joseph W. Rigney, "Effects of Graphic Presentation Combined With Graphic Review on Learning Concepts in Electrochemistry." Manuscript submitted for publication, 1977.
- Marketing News, "To Sell Products, Models Must Keep Shirts On, Research Indicates," August 12, 1977.
- Norman A. Milgram, "Verbal Context Versus Visual Compound in Paired-Associate Learning by Children," Journal of Experimental Child Psychology, 5 (1967) 597-603.
- William E. Montague and John F. Carter, "Vividness of Imagery in Recalling Connected Discourse," Journal of Educational Psychology, 64 (February, 1973), 72-75.
- Arlene H. Morgan, Philip J. McDonald, and Hugh Macdonald, "Differences in Bilateral Alpha Activity as a Function of Experimental Task, With a Note on Lateral Eye Movements and Hypnotizability," Neuropsychologia, 9 (1971), 459-469.
- Gary W. Nappe and Keith A. Wollen, "Effects of Instructions to Form Common and Bizarre Mental Images on Retention," Journal of Experimental Psychology, 100 (January, 1973).
- Robert D. Nebes, "Hemispheric Specialization in Commissurotomed Man," Psychological Bulletin, 81 (January, 1974), 1-14.
- Ulric Neisser and Nancy Kerr, "Spatial and Mnemonic Properties of Visual Images," Cognitive Psychology, 5 (1973), 138-150.
- Douglas L. Nelson and David H. Brooks, "Functional Independence of Pictures and Their Verbal Memory Codes," Journal of Experimental Psychology, 98 (April, 1973), 44-48.
- Douglas L. Nelson, Valerie S. Reed, and John R. Walling, "Pictorial Superiority Effect," Journal of Experimental Psychology: Human Learning and Memory, 2 (1976) 523-238.
- Richard E. Nisbett, Eugene Borgida, Rick Crandall, and Harvey Reed, "Popular Indiction: Information is Not Necessarily Informative," in J.S. Carroll and J.W. Payne (eds.), Cognition and Social Behavior (Hillsdale, N.J.: Erlbaum, 1976), 113-34.
- Allan Paivio, "Mental Imagery and Associative Learning and Memory," Psychological Review, 76 (May, 1969),
- Allan Paivio, Imagery and Verbal Processing (New York: Holt, Rinehart, and Winston, 1971).
- Allan Paivio and Karl Caapo, "Picture Superiority in Free Recall: Imagery or Dual Coding?" Cognitive Psychology, 5 (1973) 176-206.
- Allan Paivio, John C. Yuille, and Stephen A. Madigan, "Concreteness, Imagery, and Meaningfulness Values for 925 Nouns," Journal of Experimental Psychology, 76 (January, 1968), 1-25.
- Steven E. Palmer, "Visual Perception and World Knowledge: Notes on a Model of Sensory-cognitive Interaction," Explorations in Cognition (San Francisco: Freeman, 1975).
- G. Michael Pressley, "Mental Imagery Helps Eight-year-olds Remember What They Read," Journal of Educational Psychology, 68 (June, 1976) 355-359.
- Zenon W. Pylyshyn, "What the Mind's Eye Tells the Mind's Brain: A Critique of Mental Imagery," Psychological Bulletin, 80 (July, 1973), 1-24.
- Ronald V. Rasco, Robert D. Tennyson, and Richard C. Boutwell, "Imagery Instructions and Drawings in Learning Prose," Journal of Educational Psychology, 67 (April, 1975) 188-192.
- Michael L. Ray, "When Does Consumer Information Processing Research Actually Have Anything to Do with Consumer Information Processing?" in W.D. Perreault, Jr. (ed.), Advances in Consumer Research, Vol. IV (Atlanta: Association for Consumer Research, 1977), 372-5.
- Hayne W. Reese, "Imagery in Paired-associate Learning in Children," Journal of Experimental Child Psychology, 2 (1965), 290-296.
- Joseph W. Rigney and Kathy A. Lutz, "Effect of Graphic Analogies of Concepts in Chemistry on Learning and Attitude," Journal of Educational Psychology, 68 (June, 1976a), 305-311.
- Joseph W. Rigney and Kathy A. Lutz The Effects of Interactive Graphics Analogies on Recall of Concepts in Science (Tech. Rep. No. 79) (Los Angeles: University of Southern California, Behavioral Technology Laboratories, August, 1976b)

William D. Rohwer and Wendy J. Harris, "Media Effects on Prose Learning in Two Populations of Children," Journal of Educational Psychology, 67 (October, 1975), 651-657.

William D. Rohwer, Steve Lynch, Joel R. Levin, and Nancy Suzuki, "Pictorial and Verbal Factors in the Efficiency Learning of Paired-associates," Journal of Educational Psychology, 58 (1967), 278-284.

William D. Rohwer and Robert D. Matz (2), "Improving Aural Comprehension in White and in Black Children: Pictures Versus Print," Journal of Experimental Child Psychology, 19 (1975), 23-36.

John R. Rossiter, "Visual and Verbal Memory in Children's Product Information Utilization," in B.B. Anderson (ed.) Advances in Consumer Research, Vol. III (Cincinnati: Association for Consumer Research, 1976), 523-527.

S. Jay Samuels, "Effects of Pictures on Learning to Read, Comprehension and Attitudes," Review of Educational Research, 40 (Summer, 1970), 397-407.

Roger N. Shepard, "Recognition Memory for Words, Sentences and Pictures," Journal of Verbal Learning and Verbal Behavior, 6 (1967), 156-163.

Jack Snowman and Donald J. Cunningham, "A Comparison of Pictorial and Written Adjunct Aids in Learning from Text," Journal of Educational Psychology, 67 (April, 1975), 307-311.

Major Steadman, "How Sexy Illustrations Affect Brand Recall," Journal of Advertising Research, 9 (March, 1969), 15-19.

Amos Tversky and Daniel Kahneman, "Availability: A Heuristic for Judging Frequency and Probability," Cognitive Psychology, 5 (1973), 207-32.

Barbara W. van der Veur, "Imagery Rating of 1,000 Frequently Used Words," Journal of Educational Psychology, 67 (February, 1975), 44-56.

Melvin C. Wittrock and Sheila I. Goldberg, "Imagery and Meaningfulness in Free Recall: Word Attributes and Instructional Sets," Journal of General Psychology, 92 (1975), 137-151.

Paul M. Wortman and Phillip B. Sparling, "Acquisition and Pretention of Mnemonic Information in Long-term Memory," Journal of Experimental Psychology, 102 (January, 1974), 22-26.

Peter L. Wright, "The Cognitive Processes Mediating Acceptance of Advertising," Journal of Marketing Research, 10 (1973), 53-62.

John C. Yuille and Michael J. Catchpole, "The Role of Mental Imagery in Models of Cognition," Journal of Mental Imagery, 1 (Spring, 1977), 177-181.

John C. Yuille and Allan Paivio, "Abstractness and Recall of Connected Discourse," Journal of Experimental Psychology, 82 (December, 1969), 467-471.

VISUAL IMAGING ABILITY AS A MEDIATOR OF ADVERTISING RESPONSE¹

John R. Rossiter, University of Pennsylvania
Larry Percy, Gardner Advertising Company

Abstract

This study examines the role of visual imaging ability as a mediating variable in consumer response to advertising. A model is proposed in which visual imaging ability is hypothesized to mediate the effects of both visual and verbal stimulus content in advertising through a process which we term visual reinforcement. Data are provided in support of the model and in support of the visual reinforcement hypothesis. Visual imaging ability appears to be a powerful and hitherto overlooked mediating variable in consumer information processing.

Introduction

Advertising, in all media except radio, relies heavily on visual as well as verbal information to present the advertised product. Research on consumer response to advertising, however, has focused almost exclusively on verbal information processing; witness the plethora of advertising response studies in which verbalized beliefs, attitudes, recalled message points and so forth represent the main informational constructs. Most importantly, visual information processing has not been represented in any models of advertising response with the exception of one model developed for children by Rossiter (1976). The present study proposes a more elaborate model which examines the nature of visual information processing as it applies to advertising.

Researchers have not been blind to the importance of visual information per se. Studies of consumer spatial behavior in retail location research (for a review see Reynolds and Wells, 1977), of packaging, and of the effects of illustration size in print advertising (for a review see Hendon, 1973) are at least partially concerned with consumer response to visual information. However, these studies have employed S-R designs of an input-output type in which the nature of visual information processing has not been assessed.

The concept of information processing implies that something theoretically necessary if not directly accessible is going on inside the organism which operates on the incoming information and consequently affects the output response. Variables which tap these organismic processes constitute valid O-variables in the basic S-O-R paradigm assumed by most information processing theorists and are to be distinguished from individual difference variables such as those of the demographic, psychographic or past product usage variety which may be correlated with output responses but which have not yet been shown to affect information processing in any causative or even specifiable manner.

¹The authors are indebted to the creative staff at Gardner Advertising Company for preparing the stimuli used in the study; to the research staff at Gardner for collecting the data; and to Frank Hayson of the Wharton School for computer assistance. Professors James Bettman, Kathy Lutz and Richard Lutz of UCLA provided valuable comments on the manuscript.

O-variables which mediate verbal information include intelligence, anxiety, and sex (Hovland and Janis, 1959; McGuire, 1969). O-variables which mediate visual information are beginning to be identified in the cognitive psychology literature. An extensive review can be found in Richardson (1969).

Most of the research on individual differences in visual information processing has been concerned with the mediating effects of these O-variables on verbal learning (item recall and recognition) and motor learning (performance improvement after visualized practice). However, Richardson speculates that visually oriented O-variables may also mediate the learning of affective (attitudinal) responses. But no one has yet demonstrated visual mediation of affective learning empirically nor developed a theoretical model that would predict it. This joint theoretical and empirical endeavor is highly relevant to advertising and provides the basis for present study.

The O-variable examined in this study is visual imaging ability. Visual imaging ability may be colloquially defined as the individual tendency to "think in pictures" (Richardson, 1969). Although everyone experiences pictorial imagery at some time or other, individuals vary widely in the extent to which they employ visual imagery as a habitual mode of thinking and simultaneously in the extent to which visual imagery occurs easily and spontaneously as a covert response to external stimuli (Paivio, 1969; Richardson, 1969).

The Model

The theoretical model on which the present study is based is shown in Figure 1. The model holds a crucial implication for individual differences in visual imaging ability. It predicts that visual imaging ability will mediate the processing of both visual and verbal stimuli. The hypothesized processes are explained below for verbal stimuli and then visual stimuli.

Verbal Stimuli

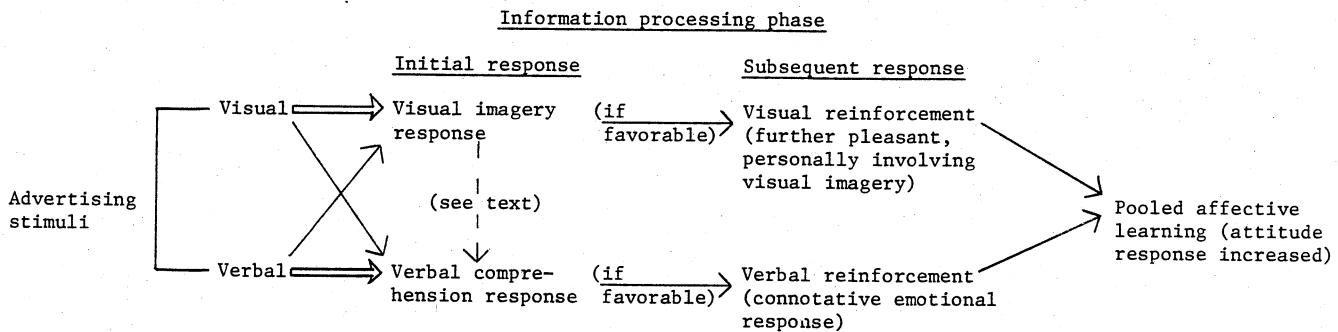
The initial and primary reaction to a verbal stimulus consists of a verbal comprehension response; that is, the verbal stimulus (word, sentence, or larger textual unit) is decoded and understood.

However, it has also been shown that verbal stimuli may simultaneously evoke visual imagery responses. Indeed, the imagery value of a verbal stimulus has consistently been shown to have a more powerful influence on verbal learning than verbal comprehension factors such as the meaningfulness of the stimulus (Paivio, 1969; 1971; Hulse, Deese and Egeth, 1975). Verbal stimuli having the highest imagery value are generally those of a more concrete nature. Concrete verbal stimuli are those which have clear, "real world" referents as opposed to vague, abstract referents. It is easier for most people, for example, to experience visual imagery in response to the word BEER than to the word BEAUTY.

The initial covert information processing response to

Figure 1

ADVERTISING RESPONSE MODEL FOR THE PRESENT STUDY¹



¹Note: The response linkages have been shown merely in graphical form. Functional relationships would follow a multistage mediational ($r_m \rightarrow s_m$) process which, for clarity, is not shown here.

verbal stimuli therefore has two hypothesized components as shown in Figure 1: a verbal comprehension response and, frequently though not always, a visual imagery response.

It is the subsequent responses, however, that are of most theoretical interest in the processing of persuasive information. The present model posits that if the verbally comprehended stimulus is recognized as a favorable one (e.g., the word GOOD vs. the word BAD) verbal reinforcement will occur; and that, likewise, if the visual imagery response to the verbal stimulus is a favorable one (e.g., a pleasant, perhaps personally involving mental picture) visual reinforcement will occur.

The verbal reinforcement phenomenon has solid empirical support. A long history of evidence beginning with a classic experiment by Staats and Staats (1957) has demonstrated that favorably evaluated words can function as verbal reinforcers. By associating objects or other words (e.g., the name of a product) with these verbal reinforcers attitudinal responses toward the object or word can be increased. Advertisers, of course, have always known this implicitly, hence their predilection for highly favorable words such as NEW or FREE in copy.

The visual reinforcement phenomenon in the processing of verbal stimuli has yet to be demonstrated. It has been shown by Paivio (1971) that the verbal learning superiority of high imagery value words occurs regardless of their connotative value; that is, "good" words are learned just as well as "bad" words providing their visual imagery values are equivalent. This at first would seem to indicate that verbal reinforcement is not involved in verbal learning but that some type of visual reinforcement, mediated through the word's evoked image, may be. However, "learning" in verbal learning and imagery studies has invariably been defined in terms of recognition or recall. This can be contrasted with the type of affective learning involved in the Staats-type experiments (and in the present study) where the dependent variable is an attitudinal or evaluative response. No one has yet investigated whether the visual imagery value of verbal stimuli relates to affective learning; that is, Paivio's finding that imagery independently increases recognition and recall can not be taken as evidence that imagery could also increase attitude because the latter type of dependent variable has not been employed in imagery studies.

Demonstration of visual reinforcement of attitudes through the use of verbal stimuli would be extremely significant for advertisers. In contrast with verbal learning researchers who try to select connotatively neutral verbal stimuli, advertisers deliberately select affectively loaded verbal stimuli to describe products. If these affectively loaded stimuli not only produce verbal reinforcement (of the GOOD type) but also visual reinforcement (of the pleasant, personally involving mental picture type) then such stimuli may constitute extra reinforcement for affectively learned responses such as attitude.

Visual Stimuli

If verbal stimuli can elicit visual imagery responses then visual stimuli should almost certainly do so. The model in Figure 1 assumes that the initial and primary reaction to a visual stimulus is some type of visual or iconic (picture-like) encoding which we have called a visual imagery response. However, as shown by the smaller arrow in the figure, a verbal reaction may also occur. Let us examine the visual reaction first.

A well known experiment by Shepard (1967) suggests that visual imagery always occurs in response to visual stimuli. Shepard's experiment is interesting because, although it was conducted in the context of experimental psychology, the stimuli employed happened to be magazine advertisements. Subjects in the experiment studied 612 illustrated magazine advertisements at their leisure. They were given a recognition test on 68 pairs of ads in which the pairs consisted of one ad from the originally studied set plus a new ad of the same type as the rest. The task was to identify which of each pair had been in the original set. On an immediate recognition test, subjects were 98.5% accurate. On a delayed test one week later subjects were still 90% accurate. Performance did not fall to the chance level (50%) until four months afterward. This normal but amazing recognition feat implies that people store visual images of every visual stimulus they pay attention to. The alternative, non-imagery explanation, that the subjects in Shepard's experiment generated verbal responses to the 612 advertisements which later aided them in the recognition test, is hardly convincing because this would imply that the subjects could accurately recall a list of 612 words in one trial, which simply cannot be done. Visual imagery therefore seems to be an inevitable

response to attended-to visual stimuli.

People may also generate verbal responses to visual stimuli just as they may generate visual responses to verbal stimuli. However, our model suggests, and the evidence on cognition indicates, that intramodal responses (i.e., an initial imagery response to visual stimuli and an initial verbal comprehension response to verbal stimuli) are the typical and dominant ones even though cross-modal responses may also occur. Dominant intramodal responses are illustrated by the larger arrows in Figure 1.

With visual stimuli, as with verbal stimuli, it is the subsequent responses that are in dispute. One school of thought acknowledges that the initial response to a visual stimulus may be a visual imagery response but that the subsequent responses are verbal. That is, the person mentally "describes" or interprets the visual stimulus in verbal terms. The verbal response then provides the basis for further effects such as recall or recognition. The best known theory in this verbal coding school is the "verbal loop hypothesis" (Glanzer and Clark, 1969). We have already argued that verbal labeling of this type would have difficulty accounting for visual recognition findings such as those of Shepard (1967). However, verbal labeling would appear to be a possibility with fewer visual stimuli than in Shepard's experiment. The verbal labeling viewpoint remains popular (Pylyshin, 1973) and stands as an alternative to the subsequent response interpretation in the present model.

The present model proposes that subsequent responses to visual stimuli may not pass through any "verbal loop" at all. It is hypothesized that the initial imagery response can serve as an internal stimulus which triggers further visual imagery. If this imagery is favorable, visual reinforcement will occur.²

The Nature of Attitudes

Several comments are necessary concerning the nature of the ultimate response in affective learning. In the present case this is an attitude (see Figure 1). Up until this ultimate response our model is quite clearly a modality specific "dual trace" model of the type advocated by Paivio (1971). Our model hypothesizes that both visual and verbal stimuli can produce verbal and visual "traces" in the form of covert stimuli and responses during information processing. However, we do not maintain that the ultimate effect, i.e., an attitude, is modality specific. Following Anderson and Bower (1974) and to some extent Pylyshin (1973) we regard the ultimate retained effect as being abstract and conceptual in nature. In retaining a response such as an attitude there is no requirement that the person also

²This view should be carefully distinguished from the imagery coding theory advanced by Bugelski (1970; 1977). Bugelski's imagery coding theory is the diametric opposite of the verbal coding theories of Glanzer and Clark and Pylyshin in that it holds that all coding is imaginal. However, Bugelski is not referring to a picture-like visual image when he uses the term "image" whereas we, in contrast, are. In this sense our view of imagery follows that of Richardson (1969) who defines visual imagery as a consciously experienced, quasi-perceptual event. We differ from Richardson, however, in that our model hypothesizes that visual imagery may also have reinforcing properties. In particular, subsequent visual imagery responses made to the initial iconic image are capable of assuming reinforcing qualities, hence our concept of visual reinforcement.

retain the original verbal or visual learning experience in toto. Just as it is possible for people to undergo an attitude change without recalling the original verbal content of the persuasive message (Greenwald, 1968; Calder, Insko and Yandell, 1974) it is similarly possible to undergo a visually mediated attitude change without recalling the original visual imagery that produced it.

Although attitudes themselves are not modality specific, attitudes learned under verbal and visual reinforcement conditions should be retrievable via appropriate cues in either modality: external stimuli similar to those under which the original learning occurred. Consequently, the product attitude response may be elicited by a visual cue, as when the product is seen again in another advertisement or in a store, or by a verbal cue, as when the product name is heard or the name read while filling out an attitude questionnaire.

Finally, the conceptualization of attitude as an automatic, involuntary, elicited response implies that visual (and verbal) reinforcement in our model operates via a classical conditioning process. All that is necessary is the contiguous association of a brand or product stimulus with a visual stimulus. The occurrence of subjectively pleasant visual imagery within the individual beyond the initial iconic encoding response then becomes a sufficient condition for visual reinforcement to occur via classical conditioning. The classical conditioning possibilities of visual imagery were first recognized by Leuba (1940) who referred to images as "conditioned sensations." This view was later elaborated by Mowrer (1960). Since 1960, with attention turning from behaviorism to the newer cognitive theories, psychologists have not addressed the potential classical (UCS) reinforcing value of visual imagery. Instead, behavioristic research on imagery has tended to follow the paradigm developed by Wolpe (1958) which focuses on the operant reinforcing value of visual imagery of a voluntary type (Cautela, 1977). Our concern, however, is with "involuntary" visual imagery and thus the classical conditioning paradigm is appropriate.

Visual Imaging Ability and Visual Reinforcement

Our model requires us to demonstrate that a visually mediated type of reinforcement occurs independently of the "standard" verbally mediated reinforcement. Verbal coding theorists could argue that all reinforcement is verbal, due in the case of visual stimuli to verbal labeling (see dashed line in Figure 1).

Examination of individual differences in visual imaging ability provides a test of the competing hypotheses. Individuals who are high in terms of visual imaging ability should be more likely to generate visual imagery to any stimulus with visual or verbal content and therefore to show greater affective learning in general because of the dual reinforcement idea. But isolation of visual and verbal content provides a more crucial test. If visualizers (this convenient shorthand will be used hereafter) show greater affective learning only in response to verbal stimuli it could be argued that they are simply generating more initial visual imagery responses to words which are then labeled and translated into verbal reinforcement. However, if it could be demonstrated that visualizers also show greater affective learning in response to visual stimuli then a clear-cut case for the validity of visual reinforcement would emerge. This is because the initial imagery response to visual stimuli is common to all individuals (cf. the Shepard experiment). Increased reinforcing effects among high visualizers would therefore make it almost impossible to avoid the conclusion that these individuals are deriving extra visual reinforcement (further spontaneously generated pleasant visual imagery) from

the same initial image because any verbal reinforcement from this image would be constant. An extreme counter-argument, that this further imagery is nonreinforcing per se but is verbally labeled and then reinforcing, seems unlikely. In any event, such an argument would not vitiate the reinforcing potentiality of visual information processing in its own right.

The Study

The test of the visual reinforcement hypothesis (and of the advertising response model in general) was devised by exposing consumers to two new brands of beer via separate print advertisements. The brand names were fictional and were equated beforehand to ensure that names per se had no effect on attitude. One brand was presented via visually oriented (strong visual) advertising: a large picture of the product with small verbal copy underneath. The other brand was presented via verbally oriented (weak visual) advertising: large verbal copy with a small picture of the product underneath. A "no picture" control condition was rejected as being atypical of beer advertisements.

Experimental conditions were further arranged so that half the consumers, in a rotated 2x2 design, received advertisements in which the verbal copy was highly concrete in nature (superlative and explicit product claims) and half received advertisements in which the verbal copy was highly abstract in nature (superlative but vague product claims).

Consumers were asked to rate their attitude toward the brand after each exposure as a measure of affective learning produced by the advertisements. Beer was selected as the product category because previous research (e.g., Allison and Uhl, 1964; Woodside, 1972) has shown that brand attitude engendered by advertising is a very good precursor of purchase behavior. Too often a haphazard selection of dependent variables such as beliefs, attitudes, or recall is made. As argued by Rossiter and Percy (in preparation) the dependent "advertising effect" variable should always be dictated by an a priori advertising response model applicable to the particular product category, consumer segment, and brand choice situation. The latter consideration (see also Wright, 1976) makes recall, for example, an unsuitable dependent variable for the beer category except, perhaps, in restaurant ordering situations where no verbal or visual cues are present. For most applications concerning beer advertising, attitude is the appropriate dependent variable.

The central issue in this study was not just whether affective learning would differ by stimulus condition but whether it would also differ predictably across individuals. A measure was therefore taken of visual imaging ability as a predictor variable. For comparison purposes, a nonvisual predictor variable, intelligence, was also included and sex differences were assessed. The main theoretical predictions, however, were: (1) visualizers would show greater affective learning than other consumers because of the greater likelihood of spontaneous occurrence of visual reinforcement from both the verbal and the visual stimulus content in advertising; (2) that this individual difference should be accentuated when the visual stimulus content of advertising favors visual imagery formation, i.e., contains strong visual emphasis; and similarly (3) when the verbal stimulus content of advertising also favors visual imagery formation, i.e., contains concrete copy. Each of these effects would provide evidence in support of the visual reinforcement component of the model.

Method

Subjects. Participants in this study were 88 adults (44 men and 44 women) recruited from a typical Mid-western suburban shopping center. All respondents were beer drinkers and an additional attempt was made to ensure a range of ages in the sample in case favored beer product attributes should vary as a function of age.

Materials and design. Four print advertisements representing the possible 2x2 combinations of visual versus verbal emphasis and concrete versus abstract copy were constructed professionally by the staff of a major national advertising agency. Visually oriented advertisements consisted of (a) a brand name at the top of the page, (b) a large picture of a mug of beer bearing the brand label and (c) concrete or abstract copy beneath. Verbally oriented advertisements consisted of (a) a brand name at the top of the page, (b) concrete or abstract copy as the main body of the advertisement and (c) a small picture of a mug of beer bearing the brand label beneath. Brand names (to serve later as the product cue) were in very large type of identical size for the visually and verbally oriented advertisements. Pictures and copy were also identical between versions except that the visually oriented advertisements used a 3:1 picture to copy size ratio whereas the verbally oriented advertisements reversed this ratio. Print size was adjusted slightly between concrete and abstract versions so that the total space allocated to copy was constant within visually and verbally oriented advertisements respectively.

The design of the study required each respondent to be exposed to both the visually oriented and verbally oriented advertising stimuli and to the concrete and abstract copy in order that visual imaging ability could be applied as a predictor on a within-subject basis. This meant that each respondent could see a visually oriented advertisement and a verbally oriented advertisement but, obviously, the brand names and copy content would have to differ.

The copy manipulation was taken care of by the concrete-abstract manipulation: each respondent simply saw the visually oriented advertisement with either concrete or abstract copy then the verbally oriented version with the complementary type of copy; this, and also the order of exposure for visually versus verbally oriented advertisements was rotated across respondents in a balanced design.

The brand name factor required the selection of two different brand names so that two different advertisement types could be rated by each respondent. To make the learning task as realistic as possible it was decided to use realistic brand names rather than abstract symbolic designations such as X or Y. The two brands of beer were introduced as new Bavarian imports (see copy below). Accordingly it was thought that Germanic sounding words would be appropriate as brand names without having immediately meaningful denotation to most consumers. The purpose would then be to select two Germanic sounding words whose connotative or affective associations were as equal and as neutral as possible so that the names alone would not influence product attitude. Nine two-syllable nouns were selected from a German dictionary. These were arranged in two randomly ordered lists. A separate pretest sample of 60 consumers (30 men and 30 women) drawn from the same population as those in the main study were then asked to rate each word on a scale of 1 to 7 in terms of how they "like [7] or dislike [1] each name as a brand name for beer." This task was explained by a short scenario which claimed that a Bavarian beer company was considering several brand names for a new import entry to the U.S. market. Two names, BAUER and LAUFER, were selected on the basis

of this pretest as being most equal and closest to the neutral affective rating position. Mean affective ratings for each of the two names were identical (3.8 for men and 3.7 for women) as were their standard deviations (2.0 and 2.1 respectively). By a coin toss, the name BAUER was used in the visually oriented advertisements and LAUFER in the verbally oriented advertisements.

The visual stimuli selected for inclusion in the advertisements consisted of identical pictures of a "foaming mug of beer." The stimuli varied only in size, location, and brand name, as described earlier. The purpose of this selection was to ensure a fairly unitary stimulus (as far as intuitive judgment is concerned since, unlike verbal stimuli, there is no standard taxonomy for visual stimuli) so that the attentional focus, which is presumably the basis of the initial imagery response, would be singular or wholistic for all respondents. Theoretical complications would arise if the visual stimulus were multidimensional. A picture of people drinking beer or a beer bottle in an environmental setting might provoke differential attention to various substimuli within the visual display. This possibility was virtually removed with the "mug shot" which at the same time represented a visual stimulus that would not be atypical of a beer advertisement. The verbal stimuli in the advertisements consisted of either concrete or abstract copy. Three beer product attributes were selected for inclusion in the copy: overall quality, taste, and price. These three attributes have been found in proprietary research to be the main determinants of beer brand attitudes. The limitation to three attributes ensured that verbal processing to generate an attitude could easily take place. The use of five or more attributes might exceed short-term memory capacity (7 ± 2) and result in differential attribute focus across respondents.

The concrete copy was designed to be superlative and explicit, of the "factual documentation" type. The text for the concrete copy was:

BAVARIA'S NUMBER 1 SELLING BEER FOR
THE LAST 10 YEARS

WINNER OF 5 OUT OF 5 TASTE TESTS IN
THE U.S. AGAINST ALL MAJOR AMERICAN
BEERS AND LEADING IMPORTS

AFFORDABLY PRICED AT \$1.79 PER
SIX-PACK OF 12 OZ. BOTTLES

The \$1.79 price was determined from the pretest by asking respondents to estimate what "affordably priced" would mean for the new Bavarian import and then computing the mean estimate.

The abstract copy was designed to present the same copy points in superlative but vague form of the more "emotional" type. The text for the abstract copy was:

BAVARIA'S FINEST BEER

GREAT TASTE

AFFORDABLY PRICED

The concrete copy is longer than the abstract copy. This is necessitated by the inclusion of sufficient information to provide an explicit referent for each attribute claim.

Procedure. Consumers participating in the study first answered some preliminary screening questions and then answered the 15-item Visualizer-Verbalizer Questionnaire (VVQ) developed by Richardson (1977) as a measure of visual imaging ability. The VVQ has been shown to have

very high convergent and discriminant validity with other measures of visual imagery, including both other self-report and also physiologically based measures of right- and left-brain processing. Test-retest reliabilities to the order of .90 have been obtained for the VVQ although its internal consistency reliability has not been established (see White, Sheehan and Ashton, 1977). The range for this test is 0 to 15. The VVQ is scored so that high scores designate individuals whose characteristic mode of thinking is exclusively visual; low scores designate individuals whose characteristic mode of thinking is exclusively verbal; with dual visual and verbal types tending toward the middle of the score distribution. This scoring method was ideal for present purposes in that visual imagery as a predictor is unconfounded by dual types of processing.

Each respondent was then shown the first advertisement which was either visually or verbally oriented and contained either concrete or abstract copy, depending on the predetermined experimental condition. After the respondent had finished looking at the advertisement, he or she was asked, "As best you can judge, based on this advertisement, how would you rate this brand?" Respondents were handed a card on which the name BAUER BEER or LAUFER BEER, respectively, was printed followed by four, 7-point bipolar rating scales. The rating scales--good-bad, inferior-superior, unpleasant-pleasant, and interesting-boring, with the end-adjectives in this order--were selected from Osgood, Suci and Tannenbaum (1957) and Fishbein and Ajzen (1975) for their high loadings on the semantic evaluation factor. The scales were scored on a 7-point -3 to +3 system with the sum (range: -12 to +12) constituting the overall affective measure of brand attitude, the dependent variable in the study. This type of multi-item index has been shown elsewhere (Fishbein & Ajzen, 1975) to be a highly valid and reliable measure of attitude. The process was then repeated with the second advertisement assigned to each respondent's experimental condition.

Finally, each respondent completed a 3-minute, timed, Number Series test. The Number Series test (Lumsden, 1959) is a highly reliable and unidimensional measure of general intelligence which correlates highly with standard I.Q. measures. The score range is 0 to 22 in terms of number correct. Its focus on numerical rather than verbal or visual stimuli makes it eminently suitable as an alternative predictor variable in the present study.

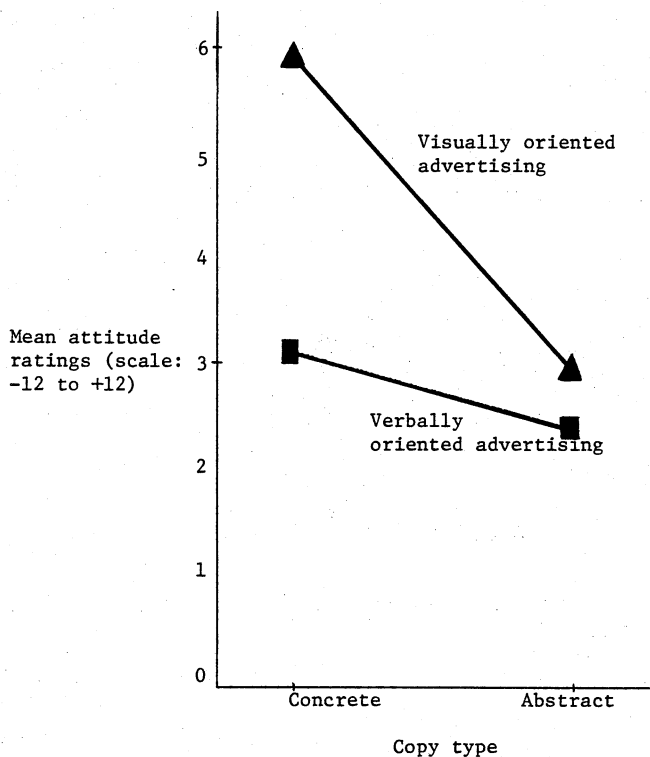
Results

The four advertisements employed in the study revealed strong stimulus effects on brand attitude ratings (Figure 2). This conclusion is warranted when it is realized that, in the pretest, the brand names alone elicited neutral attitudes. Visually oriented advertisements were more effective in inducing positive brand attitudes than verbally oriented advertisements ($p < .01$) and concrete copy was similarly more effective than abstract copy ($p < .05$). The presence of significant interaction is also evident from Figure 2, with visual emphasis and concrete copy combining to elicit the strongest attitudes. This result would be predicted by the visual reinforcement hypothesis and these overall results lend strong support to our dual reinforcement advertising model.

However, our main concern in this paper is to assess the validity of the visual reinforcement component of the model. This requires an examination of individual differences in affective learning. Accordingly, brand attitude scores were regressed against the three individual difference predictor variables, visual imaging ability (VVQ), intelligence (Number Series) and sex, plus copy type (concrete versus abstract). The correla-

Figure 2

BRAND ATTITUDE RATINGS BY STIMULUS CONDITION



tions between the predictor variables did not differ significantly from zero ($p > .05$) which eliminates the problem of multicollinearity. The results are shown in Table 1. There are three dependent variables to be considered. The first is total attitude, computed as the sum of visually and verbally induced attitudes, which provides a test of the hypothesis that visualizers will reveal greater affective learning from advertising in general. Copy type was not employed as a predictor variable for total attitude as this variable is automatically summed into the total attitude score. The second two dependent variables represent brand attitudes induced by visually and verbally oriented advertisements respectively, recalling that each respondent rated one brand under each condition.

Looking first at the total attitude results it is seen that the only significant predictor is visual imaging ability ($F = 5.32$, $p < .05$; shown to be stable in randomly split samples). Neither sex nor intelligence revealed significant mediational effects. The regression coefficient for visual imaging ability indicated that visualizers showed the greatest degree of affective learning. This finding supports the first hypothesis of the study. The bivariate correlation between visual imaging ability and total attitude was $r = .26$ ($p < .05$) which indicates that individual differences in visual imaging ability accounted for about 7% of the variance in attitude scores on an overall basis.

Looking next at the predictors of brand attitude induced by visually oriented advertising it is seen that copy type is a significant predictor ($F = 11.83$, $p < .01$; stable in randomly split samples) as is visual imaging ability ($F = 8.58$, $p < .01$; stable in randomly split samples). The sign of regression coefficient for copy type indicated that the former result was due to the previously mentioned interaction favoring visually oriented advertising with concrete copy. The sign of

TABLE 1

REGRESSIONS OF BRAND ATTITUDE RATINGS
WITH PREDICTOR VARIABLES¹

Predictor variable	Brand attitude rating		
	Total	Visual	Verbal
Copy type		11.83**	
Visual imaging ability	5.32*	8.58**	
Intelligence			
Sex			
R	.28*	.48**	.19
R ²	.08	.23	.03

¹Table entries show the significant univariate F-values (d.f. 1,86) obtained in the respective brand attitude regressions. Although not relevant to the hypotheses in this study, the multiple correlation coefficient and R² statistics are shown for summary purposes.

* $p < .05$ (1-tailed test) ** $p < .01$ (1-tailed test)

the regression coefficient for visual imaging ability indicated that, as predicted, visualizers showed significantly greater affective learning under visually oriented presentation conditions than other individuals. The bivariate correlation in this case is $r = .49$ ($p < .01$) which indicates that individual differences in visual imaging ability accounted for about 25% of the variance in visually induced attitude scores. This finding supports the second hypothesis of the study.

However, in view of the apparent interaction between visually oriented advertising and the concrete versus abstract copy dimension a further analysis was conducted by running separate regressions for the 44 respondents who received the visually oriented advertising with concrete copy and the other 44 respondents who received the visually oriented advertising with abstract copy. These results are shown in Table 2. Interestingly, the relationship between visual imaging ability and brand attitudes induced by visually oriented advertising appears to hold only for the combination of visually oriented advertising with abstract copy ($F = 12.49$, $p < .01$; stable in randomly split samples). The only significant predictor of brand attitude induced by visually oriented advertising with concrete copy is sex ($F = 7.31$, $p < .05$). However, this F value was not stable in a randomly split sample test and should be regarded with caution. Nevertheless, the regression coefficient for the sex variable indicated that women showed significantly greater affective learning from the visual-concrete combination than did men. Implications of these results for the visual reinforcement hypothesis are addressed in the discussion.

Returning to Table 1 and looking finally at the predictors of brand attitude induced by verbally oriented advertising it is seen that none of the predictors had significant univariate F-values. The absence of a significant effect for copy type indicates that visual reinforcement is not differentially triggered in visualizers by concrete copy alone. This finding counts against the third hypothesis of the study.

TABLE 2
REGRESSIONS OF VISUAL BRAND ATTITUDE RATINGS
FOR CONCRETE VERSUS ABSTRACT COPY¹

Predictor variable	Brand attitude rating	
	Visual-concrete	Visual-abstract
Visual imaging ability		12.49**
Intelligence		
Sex	7.31*	
R	.41*	.49**
R ²	.17	.24

¹Table entries again show the significant univariate F-values (d.f. 1,42) and overall regression summaries.

* $p < .05$ (1-tailed test except for the F for sex which was unpredicted and thus tested on a 2-tailed confidence region)

** $p < .01$ (1-tailed test)

Discussion

The overall results produced impressive evidence in favor of the dual reinforcement model of advertising effects. That is, across all individuals, the use of a visually oriented layout and concrete verbal copy--the two factors hypothesized to maximize visual reinforcement--generated the strongest affective learning both separately and, more powerfully, in combination.

However, the main purpose of this study was not to test the model so much as to test the validity of the visual reinforcement component of the model. The validity tests required that individuals who possess high visual imaging ability show differentially greater affective learning (1) under all stimulus conditions and especially (2) under conditions of visually oriented advertising and (3) concrete copy content. The results supported the first two of these tests but not the third. Exclusive visualizers were superior to other individuals in terms of total affective learning and in terms of affective learning induced by visually oriented advertising. But visualizers did not show any superiority to other individuals in terms of affective learning induced by concrete advertising copy.

One post hoc interpretation of this latter result is that visualizers become so preoccupied with the visual elements of advertising that they pay less attention to the copy. This may be accentuated when the copy is longer--as it was in the concrete case. The finding that visualizers showed greater affective learning with the shorter, abstract copy under the strong visual stimulus condition (Table 2) would support this "distraction" hypothesis.

The emergence of a possible sex difference in Table 2 also merits discussion.³ The sex difference occurred only in response to the strong visual stimulus-concrete

³The sex difference is described as "possible" because, as noted, the associated F-value was not stable in a split sample test.

verbal copy combination. Visual imagery is thought to be an activity which occurs, like other spatial processing, in the right brain hemisphere. Verbal comprehension, and other verbal reasoning activities, are thought to occur in the left brain hemisphere. Physiologists have found that women tend to be bilateral information processors whereas men tend to be unilateral processors (see White, Sheehan and Ashton, 1977). If so, there may be a multiplicative effect of visually and verbally induced imagery among women which leads to a stronger apparent preference for the visual-concrete combination. Women may be less "distracted" from longer advertising copy by visual stimuli in advertising because they are more able to process both types of information. This possibility has interesting implications for advertising directed to female audiences.

Also of interest was the failure of differences in intelligence to play a mediational role in the processing of persuasive information. Intelligence differences should be inversely related to persuasion when reception is constant (McGuire, 1969). However, this result has most commonly been obtained with more complex verbal messages than those employed here.

Methodological Considerations

One methodological consideration that is central to the conclusions drawn from this study concerns the possibility that the relationship between visual imaging ability and affective learning (positive attitude inducement) may be due to a "yea-saying" response bias. That is, individuals who checked the "yes" responses on the VVQ test may also have favored the positive extremes of the attitude scales used to measure affective learning.

The "yea-saying" possibility can be confidently ruled out. As noted in the Method section, the VVQ test is scored such that "yea-sayers" would fall into the middle of the total score distribution. A person answering "yes" to most or all items would emerge as a mixed imager, not a visualizer. Moreover, the VVQ test permits only dichotomous (yes-no) responses so that a tendency toward extreme responses is also precluded. The relationship between visual imaging ability and affective learning therefore remains unaffected by the measurement methods employed in the study.

A second consideration concerns the definition of "concrete" versus "abstract" verbal copy. The copy used in the concrete condition was certainly more specific and factually oriented than the copy used in the abstract condition. The idea here was to provide a singular and explicit referent for each attribute claim in the concrete copy. This is in line with the definition of concrete words as having single and explicit referents rather than multiple and often vague referents characteristic of abstract words (e.g., BEER versus BEAUTY).

However, the concrete copy required more words--regardless of the individual words' concrete versus abstract status--than the abstract copy. Our attempt to develop concrete versus abstract text was thus confounded by text length.

Nevertheless, the differential effects of the two types of text, not so much for visualizers but for the sample as a whole, remain valid. The fact that visualizers were not affected as predicted by the imagery-producing concrete text suggests that the methodology may have been in error but not necessarily the theory. In practical terms, advertisers might still see the results as generally favoring longer, factually oriented copy over shorter but vaguer superlatives.

Further Applications

The identification of individual differences in visual imaging ability may itself have limited application to advertising unless such individuals can be easily located. To be of practical use, visual imaging ability would have to be shown to be correlated with standard media demographic variables or to be prevalent in the target segment for a given brand.

Individual differences in visual imaging ability have, however, made it possible to validate the existence of the process which we term visual reinforcement--a far more general phenomenon.

The concept of visual reinforcement has provocative implications for advertising. Visual reinforcement resulting from visual imagery may explain the superior learning produced by television commercials even when television commercials and print advertisements are apparently "equated" for visual and verbal content (Grass and Wallace, 1974). The multiple visual stimuli provided in a television commercial video should produce multiple initial imagery responses and thus increase the amount of visual reinforcement experienced by the viewer.

The present study focused on positive visual reinforcement. However, our model would predict that visual imagery can have negatively reinforcing and also punishing consequences for the affective learning of brand attitudes. Advertisements for pain relievers, for example, frequently employ the negative reinforcement principle. Negative verbal reinforcement from the copy or audio may be enhanced by negative visual reinforcement from the visual elements of the picture or video. Similarly, public service announcements entailing health warnings often employ the punishment principle. These too may be enhanced by visual imagery.

Visual reinforcement therefore has many unrecognized applications in advertising and probably in other areas of marketing communication as well. Indeed, any visual or verbal stimulus which reaches the consumer may produce visual reinforcement. The present model, although tested in an advertising context, should encourage examination of visual reinforcement in other marketing communication situations.

Finally, and not to be overlooked, is that visual imaging ability represents the first visually oriented individual difference variable to be identified in consumer research. Perhaps this will speed the development of better theories of visual information processing as an obviously relevant but surprisingly neglected aspect of consumer behavior.

References

- Ralph I. Allison and Kenneth P. Uhl, "Brand Identification and Perception," Journal of Marketing Research, 1 (August 1964), 80-85.
- John R. Anderson and Gordon H. Bower, Human Associative Memory (New York: Wiley, 1974).
- B.R. Bugelski, "Words and Things and Images," American Psychologist, 25 (November 1970), 1001-1012.
- B.R. Bugelski, "Imagery and Verbal Behavior," Journal of Mental Imagery, 1 (Spring 1977), 39-52.
- Bobby J. Calder, Chester A. Insko and Barry Yandell, "The Relation of Cognitive and Memorial Processes to Persuasion in a Simulated Jury Trial," Journal of Applied Social Psychology, 4 (January-March 1974), 62-93.
- Joseph R. Cautela, "Covert Conditioning: Assumptions and Procedures," Journal of Mental Imagery, 1 (Spring 1977), 53-64.
- Martin Fishbein and Icek Ajzen, Belief, Attitude, Intention, and Behavior, (Reading, Massachusetts: Addison-Wesley, 1975).
- M. Glanzer and W.H. Clark, "Accuracy of Perceptual Recall: An Analysis of Organization," Journal of Verbal Learning and Verbal Behavior, 5 (January 1963), 289-299.
- Robert C. Grass and Wallace H. Wallace, "Advertising Communication: Print Vs. TV," Journal of Advertising Research, 14 (October 1974), 19-23.
- Anthony G. Greenwald, "On Defining Attitude and Theory," in Anthony G. Greenwald, Timothy C. Brock and Thomas M. Ostrom (Eds.), Psychological Foundations of Attitudes, (New York: Academic Press, 1968), 361-388.
- Donald Wayne Hendon, "How Mechanical Factors Affect Ad Perception," Journal of Advertising Research, 13 (August 1973), 39-45.
- Carl I. Hovland and Irving L. Janis, Personality and Persuasibility, (New Haven, Connecticut: Yale University Press, 1959).
- Stewart H. Hulse, James Deese and Howard Egeth, The Psychology of Learning (Fourth edition, New York: McGraw-Hill, 1975).
- Walter Kintsch, Learning, Memory, and Conceptual Processes, (New York: Wiley, 1970).
- C. Leuba, "Images as Conditioned Sensations," Journal of Experimental Psychology, 26 (March 1940), 345-351.
- William J. McGuire, "The Nature of Attitudes and Attitude Change," in Gardner Lindzey and Elliot Aronson (Eds.), The Handbook of Social Psychology: Volume 3, (Reading, Massachusetts: Addison-Wesley, 2nd edition, 1969), 136-314.
- O. Hobart Mowrer, Learning Theory and The Symbolic Processes, (New York: Wiley, 1960).
- Charles E. Osgood, George Suci and Percy H. Tannenbaum, The Measurement of Meaning, (Urbana, Illinois: University of Illinois Press, 1957).
- Allan Paivio, "Mental Imagery in Associative Learning and Memory," Psychological Review, 76 (May 1969), 241-263.
- Allan Paivio, Imagery and Verbal Processes, (New York: Holt, Rinehart & Winston, 1971).
- Z.W. Pylyshin, "What the Mind's Eye Tells the Mind's Brain: A Critique of Mental Imagery," Psychological Bulletin, 80 (July 1973), 1-24.
- Fred D. Reynolds and William D. Wells, Consumer Behavior (New York: McGraw-Hill, 1977).
- Alan Richardson, Mental Imagery, (New York: Springer, 1969).
- Alan Richardson, "Verbalizer-Visualizer: A Cognitive Style Dimension," Journal of Mental Imagery, 1 (Spring 1977), 96-110.
- John R. Rossiter, "Visual and Verbal Memory in Children's Product Information Utilization," in Beverlee B. Anderson (Ed.), Advances in Consumer Research: Volume 4, (Cincinnati, Ohio: Association for Consumer Research, 1976), 523-527.

John R. Rossiter and Larry Percy, Advertising Theory and Strategy (in preparation).

Carolyn K. Staats and Arthur W. Staats. "Meaning Established by Classical Conditioning," Journal of Experimental Psychology, 54 (July 1957), 74-80.

Roger N. Shepard, "Recognition Memory for Words, Sentences and Pictures," Journal of Verbal Learning and Verbal Behavior, 6 (February 1967), 156-163.

K. White, P.W. Sheehan and R. Ashton, "Imagery Assessment: A Survey of Self-Report Measures," Journal of Mental Imagery, 1 (Spring 1977), 145-170.

Arch G. Woodside, "A Shopping List Experiment of Beer Brand Images," Journal of Applied Psychology, 56 (November-December 1972), 512-513.

Peter L. Wright, "Conditional Consumer Choice Processes and Advertising Strategy," Paper presented at the American Marketing Association's Seventh Annual Attitude Research Conference, Hilton Head, South Carolina, (February 1976).

COGNITIVE RESPONSE, IMAGERY, AND SCRIPTS:
WHAT IS THE COGNITIVE BASIS OF ATTITUDE?

Bobby J. Calder, Northwestern University

Abstract

It is argued that cognitive response theory must be extended to take into account memory structure and probably visual imagery coding as well. Moreover, the notion of scripts suggests a qualitatively different view of cognitive response.

Introduction

It has become increasingly clear that to understand how attitudes are related to consumer behavior we must first understand better how attitudes themselves arise. The classical definition of attitudes as relatively enduring predispositions to respond has outlived its usefulness. Attitudes are often, perhaps usually, not enduring. They derive from what the individual is thinking about at a point in time (see, for example, Salancik, 1974, 1976; Salancik and Conway, 1975). An individual may express quite different attitudes toward the same object or behavior at different times. Not only this, equally positive expressions of attitude may be differentially related to behavior depending on the nature of their underlying thoughts (Regan and Fazio, 1977).

The purpose of this paper is to lay out what I believe are the emerging issues in understanding the cognitive basis of attitude. More questions will be raised than answered. The goal is to provide a sense of direction rather than a complete theory.

Research on the Fishbein attitude-belief model has provided a strong impetus, particularly in consumer research, to investigating the cognitive basis of attitude. The contribution of this research has been to demonstrate convincingly that attitudes are not based on a simple list of beliefs, where the beliefs are "objective" in the sense that different individuals process pretty much the same beliefs (cf. Calder, 1975). The Fishbein model is an improvement over the classic assumption originating with Hovland that attitudes are based on the rote learning of completely external (message) information. But it appears necessary to go beyond the Fishbein model to allow for an even more complex and subjective basis of attitude. It should be noted that, from this perspective, attempts merely to shore up the model by adding other variables are ill-advised.

Currently the most promising direction for research is cognitive response theory. As Lutz and Swasy (1977) have shown, cognitive response theory is essentially compatible with and a step beyond the Fishbein model. A review of cognitive response studies is provided by Wright (in press). The focus here will be on the status of the theory and how it might profitably be extended. It is argued that the theory must be elaborated to take into account the structure of human information processing. A subsequent section attempts to look beyond the present concept of cognitive response. It may be necessary to allow for alternative memory codes, such as imagery, as well as structure.

Extending Cognitive Response Theory:
Memory Structure

The new idea of cognitive response theory is that beliefs, or the thought underlying attitudes, cannot be treated as "objective." The beliefs which are pro-

cessed to yield an attitude are not just those originating in external communications; nor can they be in any sense standardized across people. Beliefs must be treated as any thought which might come to mind in a situation. The new idea of cognitive response is simply the old idea of cognitive mediation updated to a view of the person as an active information processor. Cognitive mediation is held to be better reflected in the idiosyncratic thoughts listed by a person as coming to mind after a message than by measures of message recall or standard lists of beliefs.

The cognitive response variable has stimulated research because it predicts interesting effects. For instance, Brock and his colleagues (cf. Petty, Wells, and Brock, 1977; Petty, 1977) have demonstrated that distraction can increase persuasion by inhibiting the negative thoughts which would otherwise arise with a counter-attitudinal message and that distraction can decrease persuasion by inhibiting the favorable thoughts which would otherwise accompany an attitudinally consistent message. Another illustrative effect is source credibility. In a series of studies, Sternthal (cf. Dholakia and Sternthal, in press) finds that a low credibility source produces more attitude change than a high credibility source when a person's initial attitude is already favorable. The low credibility source stimulates positive thoughts.

While such effects are intriguing, notice that there is no theoretical explanation involved beyond that of cognitive mediation. A variable's effect on persuasion is explained by whether it inhibits or stimulates cognitive response and whether positive or negative thoughts are a priori more likely. Whether positive or negative thoughts are more likely can be inferred somewhat loosely from the person's initial attitude or from the nature of the message (see Figure 1). Cognitive response is thus a mediator variable expected to proceed attitude change.

No doubt the notion of cognitive response mediation will suggest additional interesting effects. It is my opinion, however, that we must also begin to examine the nature of cognitive mediation as well as the fact of it.

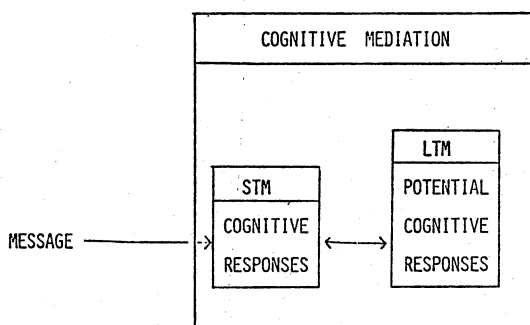
FIGURE 1
EFFECTS OF A COMMUNICATION VARIABLE ON PERSUASION

		MOST LIKELY THOUGHTS	
		POSITIVE	NEGATIVE
THOUGHTS	INHIBITS	DECREASES PERSUASION	INCREASES PERSUASION
	STIMULATES	INCREASES PERSUASION	DECREASES PERSUASION

My research has been concerned with how the structure of cognitive mediation affects persuasion. The most striking aspect of this structure is limitation. Evidence from cognitive psychology indicates that mediation is realized structurally by a limited capacity short-term memory store and a long-term memory store. The consensus model is that short-term memory holds information which is actively being processed. Long-term memory is a larger store of most, if not all, of the information that a person has ever processed. To be processed further, the information in long-term memory must be retrieved and transferred to short-term memory. The basis for this retrieval is the content of short-term memory of any point.

In terms of cognitive response theory, this model implies that information from a message, along with other incoming information, is initially represented in short-term memory as cognitive responses. These cognitive responses in turn trigger retrieval from long-term memory and the registration in short-term memory of further cognitive responses (see Figure 2). Note that any information in long-term memory is a potential cognitive response depending on what is triggered by short-term memory. But, since short-term memory is limited in its capacity, only so many cognitive responses can be represented.

FIGURE 2
THE STRUCTURE OF COGNITIVE MEDIATION

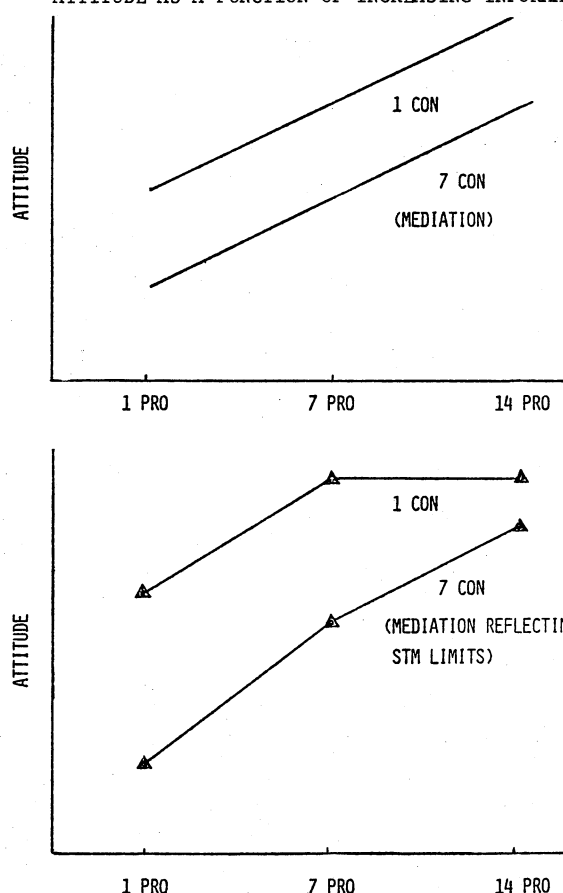


This leads to a prediction: When a message involves unfamiliar material and there is little time for rehearsal, cognitive responses to that message will be limited by the capacity of short-term memory. If the capacity of short-term memory is exceeded, additional cognitive responses cannot be represented, and thereby cannot affect attitude.

I have tested this prediction in a series of experiments which vary the amount of information in a message. Let me describe the first of these experiments. A two-sided message was used in which the pro side consisted of either one, seven, or fourteen distinct arguments and the con side of either one or seven arguments. The arguments were testimony in a jury trial. They were equated in style and content and were counterbalanced in the experimental design. After reading the message, subjects went through a thought listing procedure and indicated their overall attitude.

Without considering cognitive structure, one would expect that the number of pro cognitive responses would increase with the pro arguments and the number of con cognitive responses would increase with the number of con arguments. Attitude would thus be a linear function of the number of arguments (see Figure 3). If the number of cognitive responses is constrained by short-term memory, however, a different pattern would be expected. Beyond some point, increasing the number of arguments on one side of the message should not affect persuasion.

FIGURE 3
ATTITUDE AS A FUNCTION OF INCREASING INFORMATION



In a limited capacity short-term memory, cognitive responses favorable to the pro side must in general be represented at the expense of potential cognitive responses favorable to the con side. So the number of pro cognitive responses will increase with the number of pro arguments only to the extent that the number of con cognitive responses can be reduced. At the point where the con cognitive responses cannot be further reduced, or none are left, then further pro cognitive responses cannot be represented, and more pro arguments will have no impact.

The results of this experiment indicated that persuasion did level off with increasing number of pro arguments (see Figure 3). With one con argument, increasing the pro arguments from one to seven yielded more persuasion. The thought listing data showed that the additional pro arguments were represented at the expense of con cognitive responses. Increasing the pro arguments from seven to fourteen, however, yielded no further persuasion. The number of pro and con cognitive responses remained the same as with seven pro arguments, additional pro cognitive responses were not represented.

With seven con arguments, increasing the prosecution arguments to fourteen also yielded persuasion (see Figure 3). Changes in the number of pro arguments were reflected in changes in the number of pro cognitive responses at the expense of con cognitive responses. The reverse occurred when the number of pro arguments was decreased to one. Con cognitive responses were represented at the expense of pro cognitive responses. I believe this data demonstrates rather clearly the importance of structure to cognitive response theory.

To illustrate further the importance of structure, I would like to describe one other theoretical implication.

The more extensively information is rehearsed in short-term memory, the more likely it is to enter long-term memory. When a message is received, cognitive responses to the initial part should be rehearsed longer and therefore enter long-term memory. Cognitive responses to subsequent parts should be less likely to enter long-term memory. Now if attitudes are formed just after a message is received, two classes of cognitive responses are most likely to be represented in short-term memory. Cognitive responses from the end of the message are more likely because they should still be in short-term memory. And cognitive responses from the initial part of the message are more likely because they can be retrieved from long-term memory. Cognitive responses arising from the middle of the message should be relatively less available. This provides a more dynamic view of cognitive response. Not all cognitive responses are likely to mediate persuasion.

This implication was tested in a study conducted with Jerry Salancik. Subjects were asked to go through a long questionnaire checking religious behaviors which applied to them. Their attitudes toward religion were measured either before or after doing this. The idea was that responding to the questionnaire constituted a type of self-perception message. The effect of this message was indicated by a higher correlation between engaging in religious behaviors and expressing a religious attitude when the behavioral items were measured before rather than after attitudes. Reviewing one's behavior affected the attitude reported.

Several studies have shown this effect. What interested us was viewing the behavioral items as a message which would elicit cognitive responses. This message could be partitioned into parts by simply dividing the list of items into quarters. For each quarter we then computed a separate index of the number of religious behaviors checked. While we did not directly assess cognitive responses, we reasoned that they should reflect the items checked on each part of the list -- so that the four indices would reflect first quarter cognitive responses, second quarter cognitive responses, and so on.

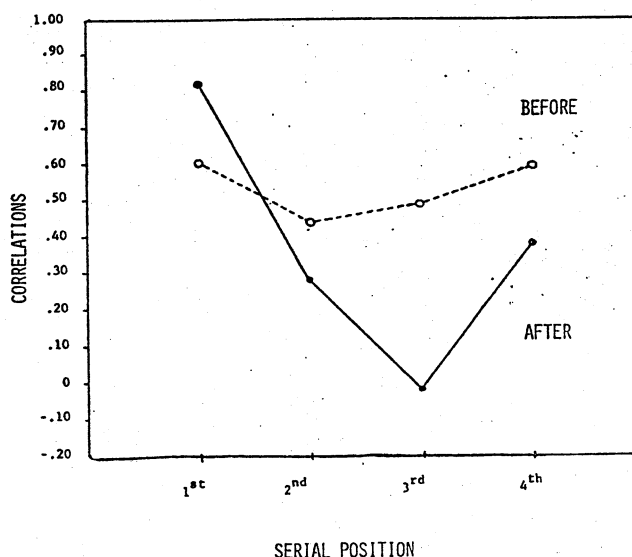
The four indices were correlated for before and after subjects with attitudes. The results are shown in Figure 4. There were no significant differences in the correlations for subjects who gave their attitudes before their behaviors. But after subjects displayed a marked serial position effect. The indices for the first quarter and the last quarter correlated most with attitudes. We conclude that more cognitive responses stimulated by the beginning and the end of the message than by the middle must have been represented in short-term memory. Since the behavior items were randomized, we can think of no other explanation besides the memorial effect.

Important methodological issues are also raised by this structural view of cognitive response theory. One is that the thought listing procedures used in cognitive response studies themselves depend on memory. Thought listing cannot be treated as automatically dumping the short-term memory contents which yield attitude.

If other activities occur after a message and before thought listing, thought listing procedures must necessarily tap long-term memory. The original contents of short-term memory will be erased by the other activities. Moreover, cognitive responses which are represented in short-term memory after a message and which do affect attitude will nonetheless be unavailable for thought listing if they have not been rehearsed enough to be placed in long-term memory. And retrieval of the thoughts which are available will be biased by the cues present at the time of thought listing. As Nisbett and Wilson (1977) have contended for other kinds of reports

on mental processes, subjects may not be able to report very well on what has affected their attitudes, and there may be good reasons why they cannot.

FIGURE 4
SERIAL POSITIONS CURVES



Extending Cognitive Response Theory: Imagery?

In a 1975 paper on the cognitive basis of Fishbein and other multi-attribute models, I pointed out that, in addition to ignoring memory structure, these models also implicitly assume that beliefs are based on linguistic verbal codes. Memory codes refer to the format in which information is stored. While it might seem that an issue as basic as memory codes should be left to the cognitive psychologist, there is a growing realization that an understanding of the cognitive basis of attitudes may require consideration of alternative memory codes as well as memory structure.

The most interesting possibility is that attitudes are based on visual imagery coding as well as verbal coding. Visual imagery is not simply pictures, or even mental pictures, though the term "quasi-pictorial" has descriptive value. It must be remembered that visual images are not objective or external, as are pictures. Imagery is a mental format for representing pictorial information. This format is thought to be the same whether it comes from visual sensation (i.e., seeing something) or from long-term memory (Hebb, 1968). Pictures are the referents of images.

If images are not mental pictures, what are they? Therein lies a controversy in cognitive psychology. Some theorists (c.g., Anderson and Bower, 1973; Clark and Chase, 1972; Polyshyn, 1973) argue against viewing visual imagery as an alternative code. The most influential argument is that of Polyshyn. He claims that all information is represented in a propositional format. Propositions are abstract logical relationships. Images and even verbal statements are surface manifestations of underlying propositions, which are themselves unavailable to consciousness. Any model of behavior must be based on propositional codes. According to this argument, visual imagery would be accorded no role in attitude theory. The experience of imagery is "epiphenomenal," that is, a mental by-product. Kosslyn and Pomerantz (1977), however, have provided a convinc-

ing rebuttal to this argument. They contend that, even if propositional coding does underlie both imagery and verbal statements, the latter have emergent properties which propositions lack. Although imagery may be constructed from propositions, the properties of visual imagery are necessary in accounting for behavior. Moreover, they doubt the necessity of even postulating underlying propositions. Contrary to Polysn, only a set of transformational rules, not a common propositional code, is necessary. These rules would specify how imagery is mapped into verbal coding.

At this point, the construct of visual imagery does seem useful in explaining many empirical findings (cf. Kosslyn and Pomerantz, 1977). A dual code (imagery and verbal coding) appears more reasonable than a propositional one. But, as this controversy points up, though the construct of imagery coding has some integrity, it is very fuzzy. Even so, it seems to me that we should begin to amend cognitive response theory to allow for imagery as well as verbal coding.

Dual-code cognitive response theory will remain pre-theoretical, however, until we can move beyond a quasi-pictorial definition of imagery. Postulating the existence of imagery coding is in itself no help in understanding attitudes. Fortunately, a recent theoretical notion called "scripts," which is attracting wide attention in social psychology, may be helpful in further developing a dual-code cognitive response theory.

Cognitive Responses as Scripts

Abelson (1976, p. 41) proposes that attitude toward an object is an "ensemble of scripts concerning that object." A script is the expectation of a sequence of events learned from direct or vicarious experience. Scripts are composed of vignettes which are the individual events, or frames, of the sequence. Vignettes generally have both a visual imagery and verbal component. Abelson uses the shorthand of "picture plus caption" to describe them. Vignettes are linked together in a coherent and causal chain to yield a script. Continuing the metaphor, the script is a cartoon strip where the separate panels combine to tell a story.

Scripts may be more or less concrete. At the most concrete level are episodic scripts. They reflect single experiences. The features of multiple experiences may be synthesized into categorical scripts. Still higher levels in which scripts are reduced completely to abstract features with no episodic character are also possible.

As an illustration, consider my experience in shopping at Marshall Field on a Saturday. What I have stored is a categorical script consisting of four vignettes: I-need-something, and the subsequent it-takes-forever-to-find-a-parking-place, followed by pushing-through-wall-to-wall-people, and finally not-being-able-to-find-a-sales-clerk. Each caption being accompanied by a suitable picture. Abelson's proposal is that it is this script, as well as any others which might come to mind, which determines the attitude I express toward the Marshall Field store. Processing this script would tend to yield a negative attitude.

The script proposal is certainly in need of further theoretical specification. I believe it suggests, however, an interesting way of relating imagery to attitudes. A product might conceivably conjure up any number of images. The script notion provides a rationale for specifying which images will be important -- those that fit together coherently to represent relevant experiences.

More generally, it seems to me that it may be useful to view cognitive responses as scripts. The usual concept

of cognitive response may be far too abstract. Cognitive responses are taken to be general thoughts about good or bad features where these thoughts are removed from any context. As scripts, cognitive responses would be firmly anchored in the stored experience of the individual. This is, of course, consistent with the major premise of cognitive response theory which is that persuasion in fundamentally self-persuasion and is, as contended by Tybout, Sternthal, and Calder (in press), initiated by self-perception.

To make the distinction between the usual view of cognitive response and script cognitive response clearer, consider an example. Suppose a woman sees a commercial for a new brand of shampoo. The usual view is that her attitude is based on thoughts about the features of the product plus any other general ideas that come to mind such as that most new cosmetic products are not very different. The script view is that her attitude is based on concrete event sequences. One cognitive response might be the categorical script ads-for-new-shampoo-products-have-caught-my-eye-before, followed by I-am-always-disappointed-when-I-try-them. Another might be the episodic script maybe-this-shampoo-would-make-my-hair-shinier, followed by my-husband-admiring-my-hair. The script view is more intensely autobiographical and far more concrete. Features and ideas do not exist in isolation but in causal chains of expectations.

The concrete character of scripts cannot be overemphasized. Nisbett, Borgida, Crandall, and Reed (1976) relate scripts to people's preference for specific information. Their example is a man deciding between a Volvo or a Saab. He might encounter all kinds of information. He might read an article in Consumer Reports based on a large sample indicating that Volvo has a better record. Or he might hear about someone's brother-in-law's who has had experience with a Volvo: "A Volvo! You've got to be kidding. My brother-in-law had a Volvo. First, that fancy fuel injection computer thing went out. 250 bucks. Next he started having trouble with the rear end. Had to replace it. Then the transmission and the clutch. Finally sold it in three years for junk." (1975, p. 129). The brother-in-law-information seems more persuasive, and there is research to support that people are biased toward such specific information. This would be expected if attitudes are based on script cognitive responses.

Conclusion

Two directions for extending cognitive response theory have been proposed. Cognitive response theory must take into account memory structure and probably visual imagery coding. Beyond this, the notion of scripts also calls attention to the need to view cognitive responses as ordered, contextual, and autobiographical. Scripts are a qualitatively different kind of memory. Given these complexities, it is not surprising that progress has been slow, and will continue to be slow, in understanding the cognitive basis of attitude.

References

- Robert P. Abelson, "Script Processing in Attitude Formation and Decision Making," in J. Carroll and J. Payne (eds.), Cognition and Social Behavior (Hillsdale, N.J.: Erlbaum, 1976).
- John Anderson and Gordon Bower, Human Associative Memory (New York: V.H. Winston & Sons, 1973).
- Bobby J. Calder, "The Cognitive Foundations of Attitudes: Some Implications for Multi-attribute Models," in M. Slinger (ed.), Advances in Consumer Research, Vol. II, 1975, 241-248.

Herbert Clark and William Chase, "On the Process of Comparing Sentences Against Pictures," Cognitive Psychology, 3 (1972), 472-517.

Ruby Dholakia and Brian Sternthal, "Highly Credible Sources: Persuasive Facilitators or Persuasive Liabilities?," Journal of Consumer Research, in press.

Donald Hebb, "Concerning Imagery," Psychological Review, 75 (1968), 466-477.

Stephen Kosslyn and James Pomerantz, "Imagery, Propositions, and the Form of Internal Representations," Cognitive Psychology, 9 (1977), 52-76.

Richard Lutz and John Swasy, "Integrating Cognitive Structure and Cognitive Response Approaches to Monitoring Communication Effects," in W. Perreault (ed.), Advances in Consumer Research, Vol. IV, 1977, 363-371.

Richard Nisbett and Timothy Wilson, "Telling More than We Know: Verbal Reports on Mental Processes," Psychological Review, 84 (1977), 231-259.

_____, Eugene Borgida, Rick Crandall, and Harvey Reed, "Popular Induction: Information is not Necessarily Informative," in J. Carroll and J. Payne (eds.), Cognition and Social Behavior (Hillsdale, N.J.: Erlbaum, 1976).

Richard E. Petty, "The Importance of Cognitive Responses in Persuasion," in W. Perreault (ed.), Advances in Consumer Research, Vol. IV, 1977, 357-362.

_____, Garry Wells, and Timothy C. Brock, "Distraction Can Enhance or Reduce Yielding to Propaganda: Thought Disruption versus Effort Justification," Journal of Personality and Social Psychology, (1977).

Zenon Polyshyn, "What the Mind's Eye Tells the Mind's Brain: A Critique of Mental Imagery," Psychological Bulletin, 80 (1973), 1-24.

Dennis Regan and R. Fazio, "On the Consistency between Attitudes and Behavior: Look to the Method of Attitude Formation," Journal of Experimental Social Psychology, 13 (1977), 28-45.

Gerald R. Salancik, "Inference of One's Attitude from Behavior Recalled Under Linguistically Manipulated Cognitive Sets," Journal of Experimental Social Psychology, 10 (1974), 415-427.

_____, "Extrinsic Attribution and the Use of Behavioral Information to Infer Attitudes," Journal of Personality and Social Psychology, 34 (1976), 1302-1312.

_____ and Mary Conway, "Attitude Inferences from Salient and Relevant Cognitive Content about Behavior," Journal of Personality and Social Psychology, 32 (1975), 829-840.

Alice M. Tybout, Brian Sternthal, and Bobby J. Calder, "An Integrative View of Cognitive Response and Self-Perception Theory," in Advances in Consumer Research, Vol. IV, in press.

Peter Wright, "Research on Ad-stimulated Thought Processes," Journal of Consumer Research, in press.

PARTICIPATION IN ORGANIZATIONAL BUYING:
SOME CONCEPTUAL AND METHODOLOGICAL PROBLEMS

Kjell Grønhaug, Norwegian School of Economics and Business Administration

Abstract

This paper focuses on problems related to participation in organizational buying decisions. Examination of previous research reveals that the most basic questions are not answered, partly due to conceptual ambiguities and methodological weaknesses.

Introduction

The present point of departure is that organizational buying may be regarded as a subset of organizational decision making.¹ An important, if not the important, element in every organization is people, that is, "organizations are social units. . ." (Etzioni, 1964, p. 3).

The frequently posed questions like who make(s) the decisions and how to reach the influentials, indicate that problems related to participation in organizational buying are important.

Based on well known, almost "classic" studies (Buckner, 1967; Harding, 1966; Walsh, 1961) leading textbooks claim that "many" people are involved in the buying decision. The widely accepted term "buying center", further indicates that not all the members of the organization need to be involved in the buying decision. Empirical research has demonstrated great variations with regard to participation both within and across organizations. However, the main impression seems to be that the number of unresolved questions still are numerous. (Zaltman, Bonoma, 1977).

Problems, Interested Parties, and Purposes

Problems

In order to proceed, let us try to pose the most basic questions. Review of the literature (Webster, Wind, 1972; Sheth, 1976) reveals that important questions are: (1) who participates, and (2) with what degree of influence? Due to the contingency perspective in organizational research (Thompson, 1967) the following question also ought to be added: (3) under what conditions?

Parties Interested

Insight related to participation in organizational buying all probably be of interest to several parties, such as:
--the buyer (in order to improve the buying system etc.)
--the marketer (in order to design effective marketing program and improve the marketing information system).
--authorities (as basis for regulations)
--researchers (as basis for further and hopefully more relevant research).
--educational institutions (as basis for educational programs).

¹Organizational buying is adopted as the appropriate term both because it is broader and more fruitful compared with terms like "industry", "bank" and so forth, and because of accumulated research on organizations assumed relevant in the context of buying. Organizational decision making has been an important research areas for decades. (March, Simon, 1958; Thompson, 1967; March, Olsen, 1976)

This list of interest groups is in no way meant to be complete, but will probably be sufficient to demonstrate variations in interests and scope.

Purposes

The various interest groups may have different purposes in mind when occupied with questions related to participation in organizational buying. Here we may make a distinction between descriptive - explanatory - predictive and prescriptive purposes. When relating the various purposes to the interest groups, we get the following matrix:

FIGURE 1

Interest Groups and Purposes

Purpose:	Interest Groups		
	Buyer	Marketer	Others
1. Descriptive			
2. Explanatory			
3. Predictive			
4. Prescriptive			

In Figure 1, some interest groups have been collapsed into "others". Obviously there will be some inter-relationships between the various purposes, such as description is made in order to explain and predict, prediction is made in order to take prescriptive actions and so forth.

From a normative point of view the buyer ought to undertake descriptive research based on relevant theories in order to take prescriptive actions. The predominant purpose for the marketer ought to be predictions as basis for prescriptive actions as it ought to be for authorities and educational institutions.

However, when making a confrontation with research conducted the following picture emerges:

--In spite of the vast amount of research related to organizational buying surprisingly few studies have focused on participation.

--Furthermore, most studies seem to have a marketing perspective in mind. The descriptive aspects, however, are predominant.

--The buyer's interests mainly taken care of in the purchasing literature on the other hand, seem to be dominated by a normative approach with almost no emphasis on the descriptive aspect.

--The all-over impression is that the various pieces of research are scattered with few (if any?) attempts to further generalizations and theory-building. (Sheth, 1976; Webster, Wind, 1972; Bonoma, Zaltman, 1977)

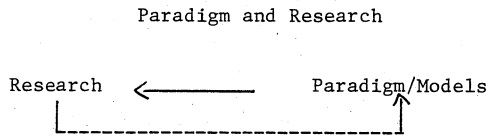
Some years ago Sheth (1972) made a distinction between four phases when assessing the research related to buying behavior with regard to the present theme, however, the research conducted has to be classified in the first

two phases, i.e. the "empirical - inductive" and probably the first part of the "formative phase" respectively.

Paradigm and Research

In order to pursue problems related to critique and explanations of the present state, and, hopefully, indicate future guidelines, let us point at:

FIGURE 2



The figure is to be interpreted as follows: Given specific problems in mind, existing "models" or "paradigms" will exert influence on the research. The broken feedback-line indicates that "paradigm"/"model" may change. However, as maintained by Kuhn (1962) and demonstrated in recent research on organizational decision-making by March and Olsen (1976), existing "models" and "paradigms" often prove to be very stable and unaffected by feed-back.

In the following we will discuss problems related to "paradigm/model" and research, respectively.

The process - paradigm

These seem to be an almost complete agreement that purchase ought to be studied in a process perspective. (Engel, 1973) However, as will be demonstrated in the following, such a perspective is not without ambiguities.

For one thing, a process has to start. Almost no attention has been given to the underlying "mechanism". In consumer buying it is often assumed that an instability such as a "felt need" or "problem recognition" represents the starting point, i.e. human needs activated by internal and external stimuli. In organizational buying this "mechanism" may prove more problematic. Where the purchase is closely related to the output such as raw materials, the "need perception" or "problem recognition" will probably be unproblematic. But when there are only loose ties between the organizational output and the potential purchase, this need in no way to be the case.

This may lead us to one often overlooked point, namely, that organizational buying, as a subset of organizational decision making, involved both cognitive and social processes. (Vroom, Yetton, 1973) Most studies have concentrated on only one aspect, mainly the cognitive, and have neglected the interaction between cognition and participation.

This perspective may furthermore be related to the starting point or "the problem recognition". In most studies on organizational buying, the assumption (also reflected in the research designs) seems to be that there exists an unambiguous stimulus situation. However, the contrary may be the case, which may require both internal and external participation. (Cyert, et. al., 1956; Hakansson, et. al., 1977)

This neglect is probably related to the almost unconscious intraorganizational perspective for most studies, which also is reflected in a commonly used definition of buying center as "... members of the organization." (Webster, Wind, 1972, p. 77) By applying an interorganizational perspective, members not included in the focal organization may participate in the purchase (Guillet de

Monthoux, 1975; Hakansson, et. al. 1977)

The buying process is often perceived as a set of stages. (McMillian, 1973) However, two problems still seem overlooked:

- (1) There may be interactions between the various stages, which also may have implications for participation and outcome,
- (2) A purchase usually involves several decisions, where the sequence between the various decisions may vary, and thus participation and outcomes need in no way to be predetermined.

In organizational buying there seems to be a firm belief in "long-lasting-relationships." (Hill, Alexander, Cross, 1975) Such Perspective, however, intuitively has an impact both on the process per se and on participation. Obviously a purchase at time t may influence a purchase at time t+1. In a study both reduction in number of decisions and participants were observed over time. (Grønhaug, 1971) This coincides with the proposition that "... cost reducing tendencies generally will function to deepen and stabilize social relations. . . ." (Emerson, 1962, p. 35)

Most (all?) studies related to organizational buying are based on the assumption that buying has to be regarded as purposeful behavior. The same observation with regard to consumer buying has been made by Sheth (1972). However, for some years, there has been some focus on artifactual decision-making, particularly as related to situations involving ambiguity and characterized by problematic preferences, unclear technology, and fluid participation. (Cohen, et. al. 1971) From the discussion above, such a perspective may also prove useful in the context of organizational buying. From the discussion so far it follows that the dominating process-paradigm involves ambiguities and has implications for participation in organizational buying only partly dealt with in previous research. Even more surprising is the picture of man emerging implicitly from the discussion above. In consumer buying the model of an active problem solver seems to be common. The neglected participant seems passive almost without influence on the buying process and outcome, just contrary to what is predominant in management thinking. (Vroom, 1976)

Part time participants

Our discussion so far has only focused on our first question, i.e. who participate(s) which naturally is related to the next two questions, degree of influence and modifying conditions.

A common observation is that people outside the purchasing department (Buckner, 1967; Grønhaug, 1977; Harding, 1966; Spekman, 1977; Walsh, 1961) may also participate in the organizational buying decision. This, however, implies that for many participants buying represents only a part-time activity, and thus the time devoted to this activity will partly depend on the total load of activities.

Degree of Influence

In spite of the importance of questions related to degree of influence, surprisingly few studies have dealt with this in the context of organizational buying. The interorganizational distribution literature, which of course includes organizational buying activities, represents an exception. (Grønhaug, 1976) When reviewing the literature, one finds the buyer described by a multitude of characteristics such as by socio-economics, attitudes, motives and personality traits. The impact of such variables on participation in organizational buying has hardly been discussed at all. However, when

reexamining the findings, variables related to insight (such as education and experience) are found positively related to degree of influence. Furthermore, insight may be viewed as one dimension of capability or a source of power. Other dimensions probably positively related to degree of influence are: political access, control over information and group support. (Pettigrew, 1973)

Conditions

With reference for the contingency perspective in organizational research, the modifying question "under which conditions" was raised. In the context of participation in organizational buying, this perspective has hardly been examined at all. This, however, does not mean that such an approach is irrelevant. From previous research it is evident that the product per se may influence participation. (Buckner, 1967; Spekman 1977; Grønhaug, 1975) The impact of individual capabilities was discussed above. Furthermore, factors related to the organization and environment have been demonstrated to affect participation. (Grønhaug, 1976, 1977; Spekman, 1977)

Research

In spite of the predominant emphasis on the buying process, very few studies, if any, really have been designed to map the processes. The interactions between process and participation have also been neglected. This somewhat limited perspective on process has naturally influenced the designs of the research conducted.

The samples studied, in terms of both underlying frames and sizes, may cause some interpretational problems. First the samples/populations studied are often restricted to one industry/area. Furthermore, many studies, in particular the more "serious" ones, are often found to be based on small samples. This does not mean that this research is inappropriate. However, at least the following seems important to consider. If certain variables are assumed to exert influence, they have at least to be present in the study. Furthermore, when going back to Figure 1, the use of research findings will often include generalizations:

FIGURE 3

Directions of Generalizations

		Organization	
		Same	New
Product	Same	1	2
	New	3	4

Different directions of generalizations from research concluded are indicated in Figure 3. Intuitively it is easier to generalize with regard to a future, purchase for the same product within the same organization than it is with regard to new products in new organizations. The degree of variation between organizations is important in this context. Frequently made generalizations based on observations from big organizations in one specific industry need not to be valid in all industries. (Buckner, 1967; Harding, 1966; Walsh, 1961) In a similar way variations across purchases may affect to what degree specific findings may be generalized, as in the Copeland's (1924) pioneering classification of

markets. Both the unit of observation and analysis will obviously influence the research. Neither isolated studies of professional buyers, such as previous studies related to "emotional" factors in buying, or the main emphasis on process seem to be adequate in order to answer the questions raised at the outset of this paper. Both information about the various actors and information related to the processes and constraints seem relevant in this context.

Research related to participation in organizational buying will of course raise measurement problems which are in no way trivial but which will be overlooked here. (March, 1955)

Concluding Comments

At the outset of this paper we pointed to three basic questions related to participation in organizational buying. These questions are in no way satisfactorily answered. In spite of this, the questions related to participation are important with relevance for both buying and selling organizations, as well as to authorities, educational institutions and so forth. How to proceed? There definitely is more than one road to Rome, and probably there does not exist one best way. However, to the present author, the following ought to be done:

- Make explicit the underlying models and paradigms in order to improve the feedback from conducted research (see Figure 2)
- Reanalyze the process-perspective in order to use it properly as basis for research, and
- Reorient thinking to include the participation-perspective.
- As noted in previous "state-of-the-art" reviews much research has been conducted. However, there definitely is a need to integrate including results from other disciplines, to advance theory construction and help the generalizations of present findings. (Sheth, 1973)
- When reviewing the literature a multitude of used concepts emerges. It is almost like new concepts are created for every problem. A better strategy would probably be to reduce the number of concepts to a few basic ones, use these intensively and concentrate our energy on the relevant questions.
- From this follows that choice of concepts and the borrowing from other disciplines should be done in a more systematic way.
- Attempts should also be done in order to reduce the time-lag between what is going on in the mother disciplines and what is borrowed.
- The various interest groups and purposes in Figure 1 may prove helpful when trying to list the research questions. Attention should also be given to when questions related to participation are relevant. The marketer's interest is related to the choices of the buyers. However, in some cases the actor has no choice, as in authority innovation - decisions, and thus questions related to participation from the marketer's side, if monopolist, are almost irrelevant. (Rogers, Shoemaker, 1971, p. 300-315) In a similar way, if there exists some type of agreement that this product--shall always be bought from. . . " then a narrow analysis of the buying center within the buyer organization will probably not prove to be very helpful for the marketer. Future research should aim at advancing our pool of relevant knowledge, and should include examination of underlying paradigms and reanalysis of previous results before adding one more empirical study.

References

- H. Buckner, How British Industry Buys, London, Hutchinson, 1967.
- M.D. Cohen, et. al., "A Garbage Can Model of Organizational Choice," Administrative Science Quarterly, Vol. 17 (March 1972), 1-25.
- M.T. Copeland, Principles of Merchandising, Chicago, A.W. Shaw Col, 1924.
- R.M. Cyert, et. al., "Observation of a Business Decision," The Journal of Business, Vol XX IX (1956), 237-248.
- R.M. Emerson, "Power-Dependence Relations" American Sociological Review, Vol. 27 (Februaruy 1962) 31-41.
- J.F. Engel and al. Consumer Behavior, N.Y., Holt, Rinehart and Winston, Inc., 1973.
- A. Etzioni, Modern Organizations, New Jersey, Prentice-Hall, Inc., 1964.
- K. Grønhaug, "Kjøpsbeslutninger pa institus jonsvare-markedet Marknadsvetande, Vol. 2 (1971) 29-34.
- K. Grønhaug, "Autonomous vs. Joint Decisions in Organizational Buying," Industrial Marketing Management, Vol. 4 (1975) 265-271.
- K. Grønhaug, "Exploring Environmental Influences in Organizational Buying," Journal of Marketing Research, Vol. XIII (August 1976), 225-229.
- K. Grønhaug, "Exploring the Impact of Organizational Variables in a Complex Buying Situation," paper presented at teh AMA-conference in Hartford, 1977.
- P.B.L. Guiellet de Monthoux, "Organizatial Mating and Industrial Marketing Conservatism," Industrial Marketing Management, Vol. 4
- M. Harding, "Who Really Makes the Purchase Decision," Industrial Marketing, Vol. 51 (September 1966) 76-81.
- R.M. Hill, et. al., Industrial Marketing, Homewood, Richard D. Irwin, 1975 (fourth edition).
- Hakansson, et. al., "Influence Tactics in Buyer - Seller Processes," Industrial Marketing Management, Vol. 5 (1977), 319-332.
- T.S. Kuhn, "The Structure of Scientific Revolutions," Chicago, The University of Chicago Press, 1962.
- J.R. McMillan, "Role Differentiation in Industrial Buying Decisions," in T.V. Greer (ed), 1973 Combined Proceedings, Chicago American Marketing Association, Series No. 35, 1973.
- J.G. March, "An Introduction to the Theory and Measurement of Influence," American Political Science, Vol. 49 (1955), 431-451.
- J.G. March and H.A. Simon, Organizations, N.Y., John Wiley and Sonc, Inc., 1958.
- J.G. March and J.P. Olsen, Ambiguity and Choice in Organizations, Oslo, Universitetsforlaget, 1976.
- A. Pettigrew, The Politics of Organizational Decision Making, London, Tvaistock 1973.
- E.M. Rogers and F.F. Schoemaker, Communication of Innovations, New York, The Free Press, 1971.
- J.N. Sheth, "The Future of Buyer Behavior Theory," in M. Venkatesan (ed) Proceedings 3rd Annual Conference 1972, Chicago, Association for Consumer Research, 1972.
- J.N. Sheth, "Recent Developments in Organizational Buying Behavior," faculty working paper, #317, University of Illinois at Urbana Champaign, 1976.
- J.N. Sheth, "A Model of Industrial Buyer Behavior," Journal of Marketing, Vol. 37 (October 1973) 50-56.
- R.E. Spekman, "The Purchasing Manager - From Order-Taker to Decision-Maker: Some Preliminary Findings, a paper presented at the AMA-Conference, Hartford 1977.
- J.D. Thompson, Organizations in Action, N.Y., McGraw-Hill, Beck Co. 1967.
- V.H. Vrom and P.W. Yetton, Leadership and Decision-Making, Pittsburgh, University of Pittsburgh Press, 1973.
- V.H. Vrom, "Leadership," in M.V. Dunnette (ed), Handbook of Industrial and Organizational Psychology, Chicago, Rand McNally Publishing Company, 1976.
- C.E. Walsh, "Reaching Those "Hidden" Buying Influences," Industrial Marketing, Vol. 46 (1961), 165-168.
- F.E. Webster, and Y. Wind, Organizational Buying Behavior, N.J., Prentice-Hall, Inc., 1972.
- G. Zaltman and T.V. Bonoma, "Organizational Buying Behavior: Hypothes and Directions," Industrial Marketing Management, Vol. 6 (1977), 53-60.

RESEARCH APPROACHES TO MULTI-PARTICIPANT DECISION PROCESSES

David T. Wilson, The Pennsylvania State University

Abstract

The study of organizational buying behavior has generated a number of concepts and models relevant to multi-person decision making. The purpose of this paper is to present some of these concepts and relate them to multi-person buying decision behavior in family decision making. A second purpose is to discuss some research approaches to multi-person buying decision making which may be relevant to the study of both organizational and family buying decisions.

Basic Concepts

The Buying Center

The concept of a buying center has been defined as "those members of the organization who interact during the decision process" (Webster and Wind, 1972, p. 77). Although the concept is intellectually appealing, most researchers have not operationalized it in their studies. This lack of use likely comes from the difficulty encountered in identifying the legitimate members of the buying center and measuring their individual influence in the decision process. Several distinct roles in the buying center have been identified: users, influencers, buyers, deciders and gatekeepers (Wind and Webster, 1972, p. 77). Although these roles are conceptually clear it is very difficult to sort out who is playing what role at a particular point in time as an individual's role may change during the decision process. Wind (1977) has neatly summarized the potential of the concept and the problem inherent in applying the concept of a buying center. He suggests that

"The relevance of the buying center concept is not limited to organizational buying, one could easily apply it to the study of consumer behavior. In this context, it is not the housewife--the traditional purchasing agent for the family--nor the entire family that should be the relevant unit of analysis. It is the family buying center which is relevant." (Wind, 1977)

Davis (1976, p. 248) points out that in most studies of family decision making "family" is usually just husband and wife and not children, relatives or friends. The concept of a family buying center would include all of the relevant influences on a purchase. Organizational buying studies should be useful to researchers studying decision making within the household as both groups of researchers must solve similar measurement problems. The cross-fertilization of ideas and approaches would benefit all.

Davis (1976) touches on many of the same research issues that Wind (1977) does. For example, the degree of influence, role versus people, number of respondents and organization of the center are issues relevant to both research areas. A reader of both papers (Davis, 1976; Wind, 1977) will be impressed by the similarities of the research problems. The concept of a multi-person buying center is useful to both groups.

The Buy-Grid

The organizational buying process can be categorized into three classes of buying situations. In each of these situations the decision process can be described by an eight-staged model (Robinson, Farris, Wind, 1967) which

is described in figure 1.

The new task situation is analogous to the extensive problem-solving situation of Howard and Sheth (1969). It is the most complex of the three buying situations and involves the largest number of decision makers and buying influences. On the other end of the scale the rebuy situation relates to automatic purchase behavior (Howard and Sheth, 1969) in which a purchase "habit" has been established. In the middle is the modified rebuy situation which occurs when the product or service purchase in a normal rebuy situation is perceived to be no longer meeting the "buyer's" needs. Product faults or competitive market forces may produce this change in the buyer's need perception.

Figure 1
The Buy-Grid Model

BUY PHASES	BUY CLASSES		
	New Task	Modified Rebuy	Straight Rebuy
1. Anticipation or Recognition of a Problem (need) and a General Solution			
2. Determination of Characteristics and Quantity of Needed Item			
3. Description of Characteristics and Quantity of Need Item			
4. Search for Qualification of Potential Sources			
5. Acquisition and Analysis of Proposals			
6. Evaluation of Proposals and Selection of Supplier(s)			
7. Selection of an Order Routine			
8. Performance Feedback and Evaluation			

The eight-stage step decision process is more suitable for organizations which document their purchasing by writing specifications than for family units who operate in a more free form environment.

Some researchers studying family decisions have used a three-phase model; problem recognition, internal and external search, and final decision (Davis and Rigaux, 1974) while others have used a four-stage model; identification of unmet needs in the form of plans, choosing the best course of action, action in the form of purchase, post-purchase evaluation (Hill, 1965). Another family decision research approach has been to examine particular aspects of a major decision (Munsinger, Weber, and Hansen, 1975) without categorizing the aspects in one of the above models.

It can be seen that organizational buying and family

buying models of the decision process have many points of similarity. Both groups of researchers are seeking to define a decision process involving multiple influences over time. A major difference is that in many instances the organizational buying process is better documented due to corporate or government regulations. Nevertheless, there are many areas where common research problems may benefit from the insights gained by one group or the others.

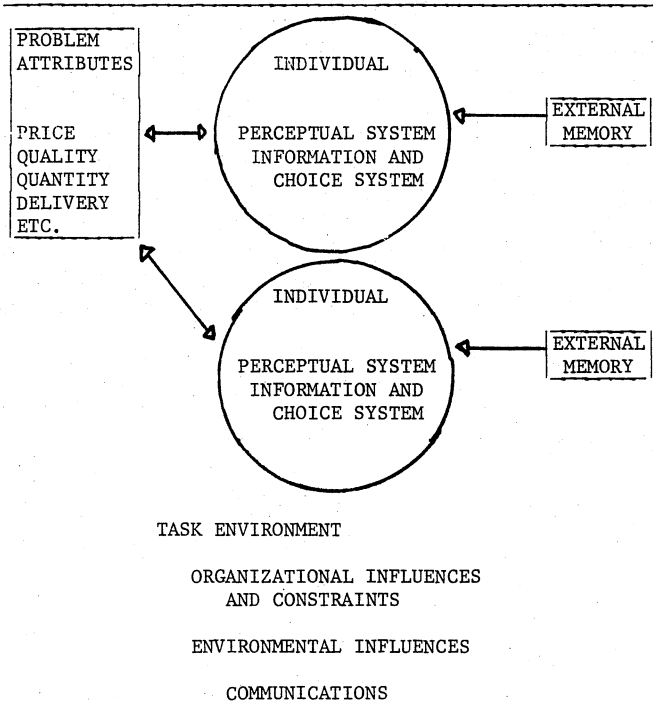
Although the buy-grid has a number of faults it does provide guidance for conceptualizing and organizing research.

A Decision Process Model

It is possible to combine the buy-grid and buying center concepts to develop a simple decision process model of the organizational buying process.

The basic structure of the model is detailed in Figure 2. The individual decision maker is the core of the model. Figure 2 depicts two individuals representing the concept of the buying center. The actual size of the buying center is a function of the buying task and the organization; in other words, its size is situation specific.

Figure 2
Structural Elements of the Decision Process Model



The term "decision process" is used as shorthand for the process that includes information-acquisition, processing activities, choice processes, and development of goals and other criteria used in choosing among alternatives (Webster and Wind, 1972: 2).

Each individual has an information processing system, the input to which is filtered through a perceptual system (Sheth, 1972, 1973). The external memory represents the notion that an organization buyer's information processing system is likely to be supported by some form of external memory ranging from a hand calculator to an on-line package of analytical programs.

The other important element in this basic structure is the conceptualization of the buying problem as a set of attributes representing the product and the vendor

(Weigand, 1966: 82; Robinson, Faris, and Wind, 1967: 40; Lehmann and O'Shaughnessy, 1974: 36). This is not a new concept, but it is an important one for this model. Attributes such as design of product, cost of application, performance life, engineering help (Weigand, 1966: 82) along with the traditional attributes of price, quality, and delivery make up the set of product-vendor attributes. This attribute set may be perceived differently by each member of the buying center in accordance with the individual's goals. Lehmann and O'Shaughnessy (1974) suggest that the importance of individual attributes varies as a function of the buying situation.

The task environment in which the individual operates likely shapes the importance of the individual attributes in total product-vendor bundle. In Figure 2, the task environment represents both environmental influences and organizational influences and constraints. Webster and Wind (1972: 40) characterize the external environment as a source of information which is input into the individual's decision process. A great deal of this information is the persuasive communications of marketers.

To review, the basic elements of the model are an individual in a task environment attempting to make a choice from a number of sets of product vendor attribute sets. This model is the most parsimonious representation of a complex process. It can be operationalized to describe simple (rebuy) or complex (new task) buying behavior.

Operationalizing The Model

It is contended that the task environment is the specifier of the operationalization of the model. The type of buying task, rebuy, modified rebuy, or new task, and the structural properties of the corporate organization greatly influence the relationship of the individuals one to another within the buying center. In addition, Newell and Simon (1972: 788) suggest that the task environment determines, to a large extent, the behavior of the individual. Their definition (1972: 55) of task environment, "refers to an environment coupled with a goal, problem or task." They note the distinction between the demands of a task environment and the psychology of the subject. This model accepts this distinction without further amplification.

Given a task environment, the four relational concepts: quasi-resolution of conflict, uncertainty avoidance, problematic search, and organizational learning (Cyert and March, 1963; Robinson, Faris, and Wind, 1967) suggest the forms that the model will assume as the buying task moves from a rebuy situation to a new task situation.

The Simple Case

In a straight rebuy situation the buyer has a clear understanding of the importance of the attributes in the situation and the possession of each attribute by each product-vendor set. A decision rule, the exact nature of which is not the concern of this paper, is applied, and a choice of vendor is made. This rebuy process can be relatively easily converted to a closed loop automatic purchasing process in which the buyer need only participate under the rule of management by expectations (Wilson and Mathews, 1971: 52) and only take over the control of the decision-making system when an abnormality exists.

It should be noted that a rebuy situation develops as the result of a new task situation. The four relational concepts have resulted in a stable rebuying situation as uncertainty has been reduced to acceptable levels and the coalition in the buying center has been satisfied.

A More Complex Case

The new task situation represents the other end of the

scale. There are many product-vendor attribute sets to be evaluated. Uncertainty is high as to the actual possession of attributes by the individual product-vendor combination. As the complexity of the task increases, a larger buying center may be created to represent groups with conflicting goals and interests.

The model can be conceived as operating in the following manner. First, responsibility for individual attributes within the bundle of product-vendor attributes is assigned to individuals within the buying center utilizing the rules of the quasi-resolution of conflict proposition. Organizational learning may be operative in that past experience may influence the current allocation of attributes to the individuals. Also, the assignment of responsibility would likely reflect the nature of the task, the skills of the individual, and other organizational variables.

A complex task situation confronts the individuals in the buying center with a likely information overload. It is suggested that the buying center seeks to reduce the complexity of the task by the easiest means possible. Problemistic search of some sort following to the principle of "information processing parsimony" (Haines, 1972) is required to reduce the number of product-vendor sets under consideration. The actual methods that buying centers use have not been addressed in the organizational buying behavior literature. Nevertheless, some speculations can be made based on studies of consumer and choice information processing behavior.

A similar decision process model could be developed for the multi-person family decision situation. Organizational variables would be replaced by variables such as stage in the family life-cycle and intergenerational influences.

The next section of the paper examines some research approaches that may be useful in the multi-person decision process.

Methodological Approaches

The complexity of multi-person decision processes in organizational and consumer buying is well documented (Davis, 1976; Wind, 1976) and has tended to retard research. Most organizational buying studies have collected data only from purchasing agents while family buying studies usually include only the husband and wife.

There is a need to develop a flexible methodology that will cut across many of the variables such as corporate size and structure. Wind (1976) has suggested that joint measurement may be one technique to quantify the effect of a relevant other person's preference on a decision maker. The amount of data that must be collected to deal with a large buying center may limit this approach but it does have promise for some decision situations.

Another approach is based upon the work in multi-attribute attitude theories. Organizational buying problems (and family buying problems) can be disaggregated into sub-problem and decision responsibility assigned to different individuals in the buying center. One way to evaluate the influence of members of a buying center may be through a version of an extended Fishbein model.

Individual members of the center would be questioned to ascertain their evaluation of the attributes salient to them. In addition, the evaluations of relevant others in the buying center, as well as their motivation to be influenced by these relevant others, would be measured. One problem will be in the individual combining to make an overall choice. It is likely that insights into this problem will come from work on noncompensatory models.

Most everyone can think of the numerous problems inherent in this approach. Nevertheless, it does offer the beginning of a method to collect and assess the influence of relevant others. There is developing a body of theory to guide researchers in developing their research approach. Hansen (1976) presents an excellent summary of these choice models.

There is a need for field work to test the viability of the above approach. For example, one way would be to analyze a number of buying situations to determine:

- (1) the individuals involved in the buying process
- (2) the attributes tacitly specified
- (3) the time-frame from initiation of project to the final choice of a supplier
- (4) an insight into decision rules for choice.

The second stage would involve individual and group interviews with members of the buying group to:

- (1) determine their views on attributes used in choice
- (2) determine the rules for the disaggregation of problems into sub-problems
- (3) determine their views on the choice processes
- (4) select several buying problems that have broad commonalities

The final stage would be the development of instrumentation to measure in field experiment:

- (1) attribute importance or saliency
- (2) attribute possession by alternative suppliers
- (3) disaggregation of the attribute set
- (4) choice rules for selecting of a supplier.

Such a study may begin to develop a methodology and instrument base founded on solid theory that would be useful for both organizational and consumer family buying studies.

Both organizational and consumer family buying involves a multi-person decision process which has many similar facets. There is much to be gained by an exchange of ideas and results by the two research streams.

References

- R. M. Cyert and J. G. March, A Behavior Theory of the Firm (Englewood Cliffs, N.J.: Prentice-Hall, 1963)
- H. L. Davis and B. P. Rigaux, "Perception of Marital Roles in Decision Processes," Journal of Consumer Research, 1(June, 1974), 51-61.
- H. L. Davis, "Decision Making Within the Household," Journal of Consumer Research, 2(March, 1976), 241-60.
- G. Haines, "Process Models of Consumer Decision Making," presented at Consumer Affairs Workshop, University of Rochester, (1971), 1972.
- F. Hansen, "Psychological Theories of Consumer Choice," Journal of Consumer Research, 3(December, 1976), 117-42.
- R. L. Hill, "Decision Making and the Family Life Cycle," in E. Shavas and G. Streib (eds.), Social Structure and the Family: Generational Relations (Englewood Cliffs, N.J.: Prentice-Hall, 1965)
- J. A. Howard and J. N. Sheth, The Theory of Buyer Behavior (New York: John Wiley and Sons, 1969)
- D. R. Lehmann and J. O'Shaughnessy, "Difference in Attribute Importance for Different Industrial Products," Journal of Marketing, (April, 1974), 36-42.

G. M. Munsinger, J. E. Weber, and R. W. Hansen, "Joint Home Purchasing Decisions by Husbands and Wives," Journal of Consumer Research, 1(March, 1975), 60-66.

A. Newell and H. A. Simon, Human Problem Solving (Englewood Cliffs, N.J.: Prentice-Hall, 1972)

P. J. Robinson, C. W. Faris, and Y. Wind, Industrial Buying and Creative Marketing (Boston: Allyn & Bacon, 1967)

J. N. Sheth, "A Model of Industrial Buyer Behavior," Journal of Marketing, (October, 1973), 50-56.

J. N. Sheth, "Buyer-Seller Interaction: A Framework," Proceedings of the Association for Consumer Research, (Winter, 1975), 382-86.

F. E. Webster, and Y. Wind, Organizational Buying Behavior (Englewood Cliffs, N.J.: Prentice-Hall, 1972)

R. E. Weigand, "Identifying Industrial Buying Responsibility," Journal of Marketing Research, (February, 1966), 81-84.

D. T. Wilson, and H. L. Mathews, "Impact of Management Information Systems Upon Purchasing Decision Making," Journal of Purchasing, (February, 1971), 48-56.

Y. Wind, "Preference of Relevant Others and Individual Choice Models," Journal of Consumer Research, 3(June, 1976), 50-57.

Y. Wind, "Organizational Buying Center: A Research Agenda," unpublished manuscript, 1977.

OBSERVATIONS OF BUYER AND SELLER TRANSACTIONS¹

Arch G. Woodside, University of South Carolina
James L. Taylor, Texas A & M University

Abstract

Wilson and Bambic's (1977) model of a multi-stage sequence of buyer-seller interactions is supported in a field observation study of forty buying centers meeting with one of three selling centers for life and health insurance. Interactions of meetings of the buying and selling centers were tape recorded. Retrospective analyses were performed separately on both buyer and selling center participants of three groups, i.e., the tape recordings of the meetings between the centers were played to the participants in separate later meetings and their verbalized thoughts and strategies were tape recorded of their prior conversations. The results support the general hypothesis that decision net models replicating mental processes in buyer-seller transactions can be developed.

Processes of Buyer-Seller Interactions

Capon, Holbrook, and Hulbert (1977) and Wilson and Bambic (1977) provide reviews of part of the buyer-seller transactions literature and these authors reach similar conclusions: most studies have been research of one stage of a multi-stage interaction process, more research is needed to identify and measure the multi-stage interaction process of buying selling behaviors, and the use of field observations of buyer-seller transactions is healthy and necessary to advance the knowledge of consumer behavior. Few field observations studies are available which include measurements of all stages of buyer-seller interactions. The research on buyers and retail-appliance salespersons reported by Willett and Pennington (1966), Pennington (1968), and Olshavsky (1973) is likely to be the most well-known literature of field research which includes measures on multi-stages in the interaction processes. This evidence does indicate that buyer-seller interactions are likely to follow an orderly process and that the successful retail-appliance salesperson is able to influence various aspects of the buyer's behavior, including the importance of product attributes, number of alternatives considered, and amount of search for information.

Additional evidence is available to support the hypotheses that a few patterns or types of buyer-seller interaction processes are likely to exist and that such patterns are likely to vary by situational variables, e.g., importance of the purchase economically or psychologically, size of the firm or family acting as the buying center. While this literature is reviewed elsewhere (Taylor, 1977; Woodside, Taylor, Pritchett, and Morgenroth, 1977), several hypotheses and findings should be noted here. Cyert, Simon, and Trow (1956) hypothesized that replicating the decision process to purchase computer equipment in a firm which they observed for several months would be possible by describing each of the activities in the process taken one step at a time (if the rules that determined when action would switch from one program step to another were specified, and if the program steps were

described in enough detail). Cyert, et al. (1956) illustrate multiple tasks performed by many persons in the buying center observed in their study and they report two common processes: communication and problem solving processes. The communication processes observed indicated the presence of gatekeepers who act as information filters to secure a large influence over the decisions reached by the decider in the buying center. Also, the search for alternative suppliers and products terminated when a satisfactory solution was discovered even though the field of possibilities had not been exhausted. "Hence, we have reason to suppose that changes in the search process or its outcome will actually have major effects on the final decision" (Cyert, et al., 1956, p. 247). For an industrial buying center purchasing computer equipment, this research suggests that search processes and information gathering processes constitute significant parts of decision making and must be incorporated in a theory of decisions if such a theory is to be adequate.

Wilson and Bambic (1977) and Bambic (1977) develop and partially test the adequacy of a five stage decision process to describe buyer-seller interactions in industrial marketing situations. Their dyadic process model is concerned basically with the development of long-term buyer-seller relationships rather than "one-shot" selling situations. The model begins with an initial meeting between the buyer and seller and moves through a number of stages which presumably take place over time and a number of meetings of the dyad:

1. Source legitimization - the buyer establishes whether the salesperson is a representative of a qualified source of supply
2. Problem identification and information exchange - the salesperson identified problem and seeks information from the buyer
3. Attribute delineation - buyer-seller determine product attribute sets which may solve problems
4. Attribute value negotiation - negotiation of the attribute set and the exchange rate for each
5. Relationship maintenance - the dyad members maintain and build upon their business relationships; perhaps the relationship becomes personal.

The duration of the stages overlap each other as a number of activities may be carried on simultaneously among buyer-seller interactions. However, the requirements of the basic stages must be met in order to move to the advanced stages (steps 4 and 5) of the model (Wilson and Bambic, 1977).

Additional research evidence is reported here in an attempt to support the general hypotheses that a few models of buyer-seller interactions can be developed to describe the stages in the decision processes.

¹Financial support of the Division of Research, College of Business Administration, University of South Carolina, and agreement to participate of company officials in a national life insurance firm (anonymous in this study) which made this research possible are appreciated.

Field Observations of Buyer-Seller Interactions for Life Insurance

Buyer-seller interactions in the field of life insurance was selected for analysis because of previous survey and laboratory research in this field (Evans, 1963; Gadel,

1964; Busch and Wilson, 1976), the large financial investment for life insurance by the buyer, and the multiple meetings and generally complex nature of the interactions. The general purpose of the research is to build upon prior life insurance research of buyer-seller interactions, and the work of Cyert, et al. (1956), Olshavsky (1976), and others (Woodside and Fleck, 1977; Bettman, 1974; Tucker, 1967) who advocate the need to know the decision process of each member involved with the buying center with respect to the task at hand.

The research approach is an application of inductive model building instead of the more widely used deductive method of theory construction of buyer-seller interactions. Olshavsky (1976), Newell and Simon (1972), and Cyert, et al. (1956) advocate an inductive approach as an initial step to theory development through the development of descriptions and programs of decision processes. This inductive approach begins with observing buyer-seller interactions and attempts to construct anatomies of decision processes, while the deductive approach begins with insight and attempts to test conceptualizations to behavior later observed. Sheth's (1976) model of buyer-seller interaction is an example of the deductive approach to theory construction. The work by Tucker (1967) is an early example of the inductive approach.

Method

The national, regional, and local headquarters of an insurance company agreed to participate in an observation study of selling and buying centers in Columbia, South Carolina. Forty prospective life and health insurance buying centers agreed to cooperate and permit tape recordings of the meetings with members of the selling centers (two buying centers contacted refused to have the meetings tape recorded). In 17 (43%) buying centers, two persons met with the salesperson (s) during the encounter; while in 23 (57%) encounters only one person met with the salesperson (s). Thus, a total of 57 buyers actually met with salespersons. Ten buying centers met with the salesperson two separate times, while 30 buying centers met with the salesperson one time each.

Buying centers were selected for the study on the basis of convenience by the three salespersons in the regular course of their selling duties. The buying centers participating appeared to the local sales manager and researchers to be representative of the type of client normally sought by the participating salespersons. Twenty-two (55%) of the buying centers were contacted initially by the salesperson via telephone, six (15%) by direct mail, and twelve (30%) by "cold call." A "cold call" refers to a salesperson's unannounced visit to a prospect's place of business or home.

The three salespersons selling experience ranged from 1 to 5 years with the cooperating company and their ages from 28-34. None of the salespersons sold insurance for any other firm. Two salespersons were white males and one was a black male. All had at least some college training ranging from a degree in insurance at a local technical school to a Masters in Business Administration.

Data Collection Procedure

Two types of data were collected. First, encounters between salespersons and prospects were tape recorded, and second, postpurchase in-depth interviews were conducted with members of three selling and buying centers. The selection of centers from in-depth interviews was done randomly among the 40 buying centers.

The in-depth interviews following the meetings of the buying-selling centers permitted a "retrospective analysis" (Olshavsky, 1976) of persons in each center, i.e., the tape recording was played back to the buying and

selling centers separately and the participants were requested to comment on the meanings and strategies behind their own statements in the interactions.

The conversations between salespersons and buyers were recorded by the researchers while accompanying the salespersons when calls were made on potential clients. In 32 (80%) of the initial meetings the buyer(s) did not know that a researcher would be accompanying the salesperson before arrival. This was a result of the difficulty in synchronizing work schedules. Additionally, the salespersons believed that permission to record the transaction was better obtained when the buyer(s) could meet and evaluate the researcher in person.

The participating buying centers were told by the researchers that the researchers were interested in determining if people who meet for the first time or who have known each other for a short time communicate differently than persons who have known each other for a long while.

All three buying centers contacted for the retrospective analysis agreed to meet with the researchers for separate interviews. Members of these three buying centers were first asked a series of open-ended questions, e.g., "When did you first come into contact with (the salesperson)?" "What were your feelings when the salesperson first contacted you?" Answers to these open-ended questions were tape recorded. Then, small segments of the original transaction was played back to the buying center members. Individuals were questioned as to their thoughts, feelings, goals, and strategies after listening to each of these tape segments of their previous conversation with the salesperson. Buying center members were also informed to stop the recording at any point that they desired to make a comment.

Buying center members who were not present when the salesperson-prospect encounter took place (e.g., neighbors, friends, family), were identified by asking subjects what other persons they had discussed their purchase decision with. These buying center members were then contacted and interviewed by the researchers. Two such persons were interviewed by telephone.

Salespersons involved in the chosen interactions were interviewed in a similar manner. The salespersons were first asked (1) how they became aware of the prospect(s), (2) how they contacted the prospect(s), and (3) what their initial goals and strategies were when making first contact. Then, the original transaction was played back after instructing the salesperson to stop the recording at any point to comment on goals, strategies, thoughts and feelings. The recorder was stopped after playing small segments to probe for these mental processes.

Analysis

Protocol analysis and correlational analysis were utilized to analyze the data. Verbal and graphic process models were developed for three buying centers and for three selling centers using protocol analysis.

Protocol analysis for the purpose of this study includes the modification suggested by Olshavsky (1976). Generally, a protocol is a transcript of the verbalized thoughts and actions of a subject when the subject has been instructed to think or problem solve aloud (Alexis, Haines and Simon, 1968). A protocol in this study is a transcript of verbalized thoughts and actions obtained by asking subjects to relive their mental processes through recorded conversations. In either approach, the transcript is a record of the subject's thought processes while engaged in making a decision.

The theory of human thinking and problem solving from which protocol analysis sprang postulates that the human operates as an information processing system (Newell,

Shaw and Simon, 1958). The theory posits that there exists: (1) a memory, (2) search and selection procedures, and (3) a set of rules or criteria which guide the decision making process.

The tasks in the protocol analysis were to reveal the individual's mental processes through in-depth interview and to reproduce the process in the form of a graphical decision net. Similar applications of the technique to marketing problems have been reported (Cyert and March, 1963; Morgenroth, 1964; Howard and Morgenroth, 1968; Woodside and Fleck, 1977).

Decision processes are described in terms of inputs, information processing decision rules and outputs (Howard, Hulber, and Farley, 1975). The models in this study are represented graphically as a set of decision nodes with arcs (lines connecting pairs of nodes). Each decision node represents a test on a particular decision cue (such as, should I meet with this salesperson?) and the arcs represent the processing sequence taken (yes or no), depending on the values of the cues.

Generalizations and hypotheses about buyer-seller behavior in insurance exchanges were inferred from the decision models. The findings of the protocol analysis are not meant to be conclusive, but may provide a platform to the development of theory and future research. These decision models of buyers and sellers interacting, along with similar models, may serve as valuable data points from which more universal generalizations about consumer and salesperson decision-making in naturalistic environments can be made.

The findings which follow refer only to the interactions and decision processes of the buying and selling centers included in the retrospective analysis.

Process Model Showing Buyer-Seller Interaction Effects for Buying and Selling Centers A

Buying center A consisted of one prospect, Kathy Berryman, and two other persons who participated in this prospect's decision-making process. Only Kathy Berryman interacted with the salesperson representative from the selling center. Selling center A consisted of the salesperson Jed Sutton, and the agency manager, Stan Trotter. Only Jed Sutton interacted with the buyer. Stan Trotter's influence on the transaction was indirect through the advice provided Jed as to the need of being flexible in the sales encounter. This advice was offered at a sales meeting one week prior to Jed's meeting with Kathy.

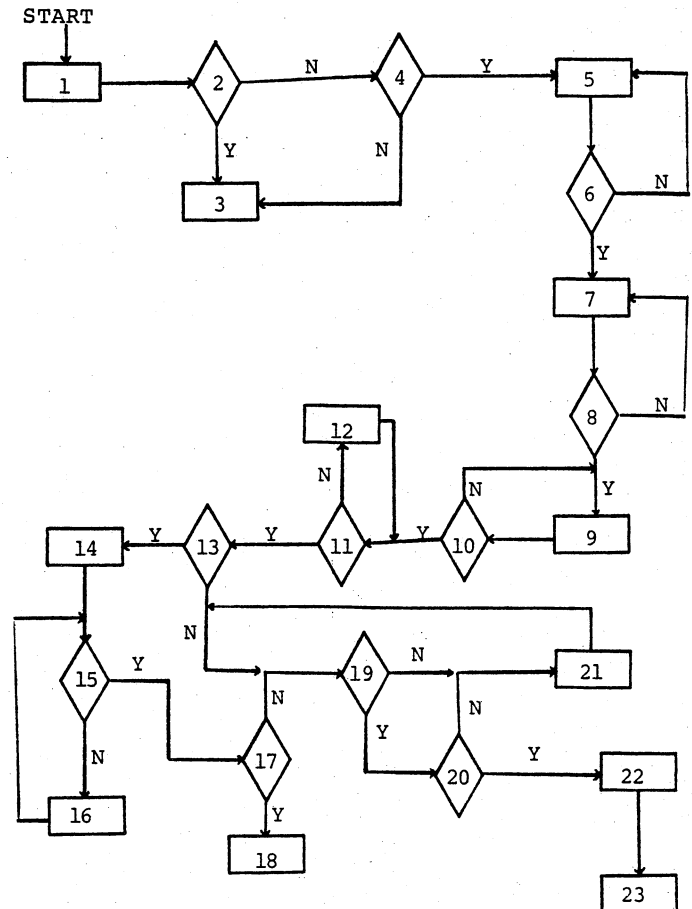
Two separate one-hour in-depth interviews were conducted with both the salesperson and the buyer. Both interviews with Kathy Berryman were conducted in her home; one interview with Jed Sutton was in his home and another in his business office. The two members of the buying center who did not meet with the salesperson were contacted and interviewed by telephone. Conversations with these latter buying center members lasted approximately 10-15 minutes.

Verbal and graphical process models for both Jed Sutton and Kathy Berryman were developed from the retrospective analyses. The models first were developed from the interview data and then presented to the salesperson and buyer for possible revision. Both parties indicated necessary revisions and then, that the models accurately represented the content of their mental activity during the sales encounter.

The verbal process models for the salesperson and buyer are included in Table 1. The words expressed in the verbal models are generally those of the participants. Some editing was required, however, for purposes of brevity and clarity. The graphical models of salesperson and buyer are shown in Figures 1 and 2 respectively.

FIGURE 1

Jed Sutton's Insurance Selling Decision Model



Key to Figure 1

Dictionary

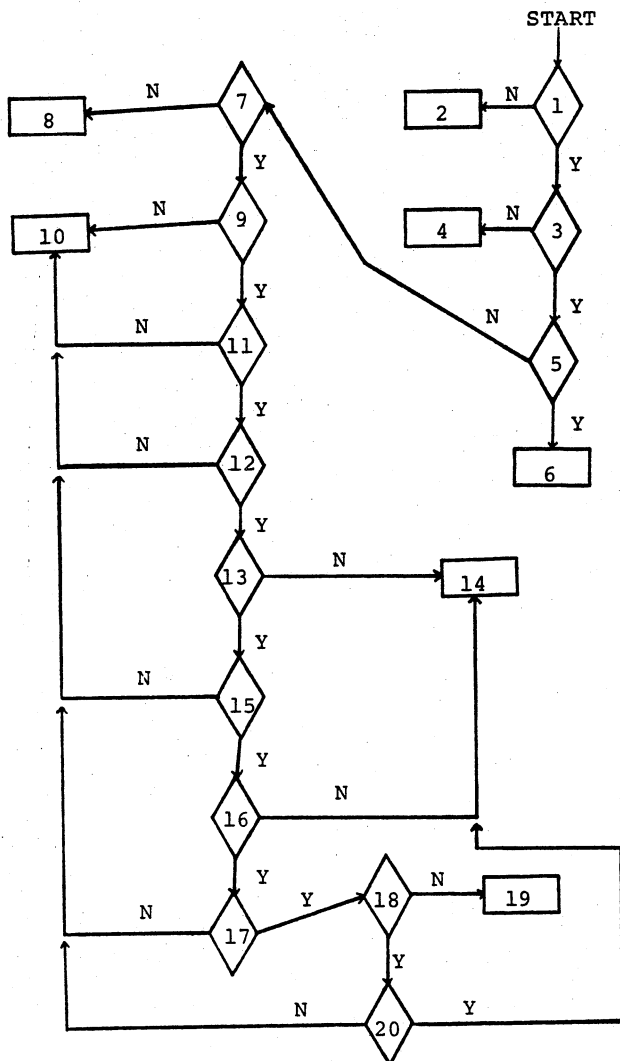
1. Review prospect's present coverage.
2. Does prospect's present mortgage policy meet her needs?
3. Do not try to make a sale with this prospect.
4. Will prospect meet with me?
5. Establish rapport with prospect by emphasizing similarities between us and by references to familiar others.
6. Have I established sufficient rapport with prospect?
7. Discuss reasons prospect does not need mortgage insurance (such as prospect's age and marital status).
8. Have I convinced prospect that she does not need mortgage insurance?
9. Establish a need for whole life coverage by discussing cash value and paid-up features.
10. Have I established a need for converting the mortgage policy to whole life coverage?
11. Is prospect convinced that a whole life policy will supplement other policies presently held?
12. Discuss endowment features of present policies.
13. Does prospect seem to be more interested in disability income protection?
14. Revise goal to sell disability income policy.
15. Does prospect understand what social security benefits she would receive if disabled?

16. Explain social security benefits.
17. Is prospect still interested in disability income protection after learning of social security benefits?
18. Attempt to close sale on a disability income policy.
19. Have I established a need for converting to at least \$6,000 in whole life coverage?
20. Will prospect convert to whole life if payments do not exceed \$20.
21. Reiterate merits of converting to whole life.
22. Close sale of whole life policy by completing application and getting a check from prospect.
23. Make attempt to develop a long term relationship as customer and agent.

Y = Yes
N = No

FIGURE 2

Kathy Berryman's Insurance Purchase Decision Model



Key to Figure 2

Dictionary

1. Do I need to review present coverage?
2. Reject salesperson's offer to meet.
3. Is this salesperson acceptable to me?
4. Discontinue interaction.
5. Do I need mortgage insurance?
6. Keep present mortgage insurance?
7. Should I consider converting the mortgage policy to a whole life policy?
8. Terminate mortgage policy.
9. Does whole life policy provide sufficient cash value at age 65?
10. Reject whole life policy.
11. Is paid-up insurance a feature of the whole life policy?
12. Is the whole life policy compatible with my other coverage?
13. Should I investigate disability income protection?
14. Convert mortgage policy to whole life.
15. Does brother-in-law recommend converting?
16. Does brother-in-law recommend exploring merits of disability income protection before deciding on conversion?
17. Does lawyer friend recommend converting to whole life?
18. Do social security benefits provide adequate disability protection?
19. Evaluate merits of disability income policy?
20. Will \$20 per month purchase adequate whole life protection?

Y = Yes
N = No

The Salesperson

Jed Sutton is a 30-year-old white male who is married and has a two-year-old son. Jed has been a salesperson with the cooperating company for one year.

Jed graduated from the University of South Carolina with a bachelor of science degree in electrical engineering and economics. After graduation, he served four and a half years in the U.S. Navy as a submarine officer. The rank of Navy Lieutenant was obtained while on active duty. Jed is still a member of the Navy Reserve.

After completing military service, Jed returned to the University of South Carolina and earned a master's degree in business administration (MBA). He then worked for Proctor and Gamble as a production manager for two years. Jed left Proctor and Gamble to go to work as a salesperson for the cooperating company. He also serves as an ACE consultant for the Small Business Administration.

The Buyer

Kathy Berryman is an unmarried white female who is 50 years old. She works as a nurse for the American Red Cross and attends the University of South Carolina on a part-time basis. She is working toward a master's degree in public health service at the University. Kathy owns her own home which is valued at approximately \$46,000.

Preexchange Activity

Kathy Berryman had previously purchased a mortgage protection (term insurance) policy from the cooperating company. The salesperson who had sold her the policy had left the company and Jed Sutton had been assigned to handle her case. Jed had reviewed her coverage and

TABLE 1

Verbal Process Model for Buying and Selling Centers A

Salesperson's Decision Process

Prospect's Decision Process

Prior to Meeting

I review the prospect's coverage and find that she has a \$30,000 term insurance policy. Her policy builds no cash value over time. She does not need \$30,000 worth of protection because she is 50 years old, has no dependents and has two other endowment policies. My goal is to convert the term policy to a \$10,000 whole life policy. My strategy is to show the prospect that she is spending \$20 a month when she has no one to protect, therefore she should do something for herself.

The salesperson calls to inform me that he has been assigned by his company to review my coverage. The agent who sold me my old policy has left the company. The salesperson recommends we get together and talk about my present coverage. I wonder whether my present coverage should be reviewed. A long time has passed since I last looked at my coverage. I may need to revise my policies. I consent to meet with the salesperson.

First Meeting

0-5 Minutes

I arrive at Kathy's home as scheduled. My first goal is to build rapport with the prospect. My strategy for doing this calls for talking about things of interest to Kathy. We talk about Columbia and the local university. I discover a common ground between us. Kathy knows and likes a professor that I once worked for. I make several references to this fact.

The salesperson arrives as scheduled. My first goal is to find out more about him. I ask how long he has lived in Columbia. I find that he graduated from the local university where I attend as a part-time student. I ask if he has ever met a particular professor that I know and like. I am surprised to discover that the salesperson once worked for this professor.

5-7 Minutes

I ask the prospect if anyone is dependent on her. My strategy is to establish that no one would suffer an economic loss at her death. Kathy needs only sufficient insurance to cover burial expense, since the value of her home far exceeds the remaining mortgage. Her house could be sold to pay off the mortgage. Thus, \$30,000 worth of term insurance is too much.

Jed asks if anyone is dependent on me. I admit that no one is. I wonder why he brought this point up. He probably wonders why I have insurance with no one to protect. I remember that my mortgage company recommended, but did not require, that I purchase a mortgage protection policy when I bought the house.

7-10 Minutes

I review the prospect's present policy with her. I show how the value of her present policy has decreased to \$26,400 and how the value will continue to decrease to \$8,400 by age 65. Her payments remain the same. My strategy is to establish a need for revising her coverage.

The salesperson tells me how the value of my policy is decreasing while my payments remain the same. He draws a graph to show this. My goal here is to comprehend what this means. I want more information.

10-15 Minutes

My strategy is to establish an economic motive for converting to a \$10,000 whole life policy. I explain the accumulation of cash value and the net costs (premiums minus cash value) associated with the whole life policy. I try to show that, at age 65, Kathy would have some accumulated cash value and would have the option of (1) continuing the policy, (2) taking a paid-up policy, or (3) taking the cash value.

I am getting a lot of information. I understand the cash value part but am not sure I comprehend the different options completely. I explain that I purchased my present policy to cover part of my mortgage. I ask the salesperson to explain the options again.

15-18 Minutes

Kathy is confused on the options. I attempt to clarify what choice she can make at age 65. I emphasize the flexibility she would have with the whole life policy.

I am beginning to understand the different options I would have with a whole life policy. I wonder whether my other insurance policies should be taken into account. I ask Jed about this.

19-21 Minutes

Oh, oh! I knew this was coming. I know she has two policies but am not sure of the face amounts or types. I must be flexible here. I ask about the face amounts and when they endow. The face amounts are \$2,000 and \$3,000 and both endow at age 65. I am in luck. Since both endow at age 65, her coverage stops at that time. My strategy now is to show her that in order to have coverage extending beyond age 65 she should convert to

Jed tells me he assumed I had other coverage. He asks about the face values and when they endow. I tell him from memory because I am not sure where they are. I ask if it might not be a good idea to look at these two policies before deciding.

the whole life policy recommended.

21-24 Minutes

I attempt to educate Kathy on endowment policies. I see no advantage of having to look up the policies since I know what they are designed for -- retirement.

Jed tells me endowment policies are usually designed for retirement, thus he is sure my policies endow at age 65. Since I am not sure where the policies are, I will drop the matter of looking at them.

24-27 Minutes

I inform Kathy that the whole life policy would take up protection where the others leave off. My strategy is to get back on track (i.e., talking about the whole life policy).

I was not aware that my protection ends at age 65. But I do have protection until that time.

28-32 Minutes

Kathy has some coverage until age 65. Thus, I revise my goal. I try to convert to some amount less than \$10,000. I suggest she convert to whatever amount would keep her payments the same as now (approximately \$20).

Jed tells me I can convert to an amount less than \$10,000 and keep my payments the same. If I change the policy, this is what I would probably do. But, I'm not ready to decide yet. I want to think about it for a while.

32-33 Minutes

I solicit a commitment from Kathy by asking which way she wants to go -- \$10,000 or a lesser amount.

I don't know what I want to do. I want more time. I know the value of my house is greater than \$26,400. The reason I bought the policy in the first place was to pay off most of the mortgage. I ask Jed if the policy would not still do this even though my house has increased in value. I think I am asking this to get more time to think.

33-35 Minutes

Kathy's not ready to commit herself. She digresses to talking about her mortgage. I want to get her back on track. I point out that, since the value of her house exceeds the mortgage, the house could be sold to pay off the mortgage. My strategy is to establish that the mortgage policy is not needed since no one depends on her.

I am beginning to believe that I should not have bought a mortgage protection policy in the first place. I wonder why my old agent sold it to me. I guess he sold it to me simply because that is what I asked for.

35-37 Minutes

I try again to close. I ask whether she would prefer to go with the \$10,000 conversion or with the \$20 a month. I recommend the \$10,000 conversion.

I'm being asked what I want to do. I want some time to think about this. I tell Jed I want to think about it and talk it over with a friend of mine before deciding.

37-40 Minutes

Two close attempts have failed. She appears to be leaning toward conversion but uncertain what to do. I try again to get a commitment by showing how easy it is to convert. I stress administrative factors of converting that make it desirable to act now. My strategy is to establish that by acting now she will be able to keep her payment date the same. But, to do this I must send her new application in soon.

Jed tells me that I need to get the application in as soon as possible to keep my payment date on the 10th. This is not important to me. I just want time to think all this through before deciding. If I do not need so much mortgage protection, maybe I should investigate something like disability income protection. I ask Jed about this

40-43 Minutes

Oh, boy! Kathy has shifted to disability income. This may be a potential sale, but for now I want to keep discussion on the whole life alternative. I mention the "waiver of premium" option on the whole life policy. This pays policy premiums if the insured becomes disabled. My strategy is to get back on track.

Jed tells me about a "waiver of premium" option. This isn't exactly what I had in mind but it sounds like a good idea. I still want more time to think about all this, however. I change the subject by asking Jed where his home is.

43-45 Minutes

It's obvious Kathy is not going to purchase tonight. My strategy at this point is to ensure that I get another appointment with her.

Jed asks when we can get together again. We discuss several dates and set up a time. I ensure that I will have enough time to talk to several friends before we meet again.
(Remainder of first and second meetings are not described here.)

45-50 Minutes

Before leaving, my strategy is to reiterate all my main selling points because I know that I will be competing with some unknown third party. I accomplish this.

Jed summarizes the merits of converting. I think of these things after Jed leaves. I'm trying to assure myself that the change is necessary and a wise move.

50-52 Minutes

We engage in a few minutes of "small talk" and I leave.

We talk about school for a few minutes and Jed leaves.

Between the First and Second Meetings

Things are out of my hands now. If the "friend" advises Kathy not to convert, my strategy will be to get both Kathy and her friend together so my side of the story can be heard. I hope this does not become necessary.

I call my brother-in-law in Greenville. I ask him what he thinks I should do. He advises me to look into a disability income policy. He thinks that it would be better for me to have some income protection in case I became disabled. He adds that he does not believe that any agent would fail to advise me properly but that I should ask about the disability plan to be sure I had considered all alternatives. I call another friend. He is not much help. He just tells me that the agent would probably not mislead me. I want him to tell me whether I would be making a wise move by converting but he avoids being explicit about this. I decide to pursue information on disability income protection. This is my goal for the second meeting.

Second Meeting

0-4 Minutes

Kathy arrives at my office. She asks about disability income protection. I change my goal to one of selling her a disability income policy. My strategy is to educate Kathy on disability income protection. I explain the 39, 60, 90 and 120 day elimination period plans.

I arrive as scheduled at Jed's office. I tell Jed about talking to my brother-in-law. I ask if my policy can be converted to a disability income policy. Jed says no. But, he explains the different elimination period plans. The smaller the elimination period the more expensive the plan. I think of this as being much like deductible automobile insurance.

4-9 Minutes

I explain how Kathy would receive \$450 a month after 90 days with the 90-day elimination plan. I mention she would get less per month for a partial disability. I attempt to show that this plan would augment social security benefits. I find Kathy is unaware of the social security benefits.

Jed provides information on a specific plan. He mentions social security benefits. I have to work 12 more years for social security. No??? Jed says I would have disability benefits coming after one year. I seek more information on this.

9-15 Minutes

I compute that Kathy would receive \$422 per month after one year if she became disabled. I tell her this. I then simply let her talk for awhile to determine what she thinks of this.

I'm surprised to find out what benefits I could receive through social security. I'm not sure that I need disability income in light of this. Converting to a whole life policy may be a better idea. I ask Jed what's involved in converting my term policy.

15-20 Minutes

We are back to conversion of the term policy to a whole life policy. How did this happen? Kathy appears to have had a change of heart concerning disability insurance. I tell her what needs to be done. My goal changes to getting her to convert to whole life. My strategy calls for just being flexible.

I have decided to convert my present policy to whole life but only to such an amount that my monthly premiums remain the same as now. I decide this after considering the social security benefits.

20-26 Minutes

I reiterate the merits of converting. I believe we are at a compromise point in which she will buy, but I need to find out what amount she desires. I try to probe for the amount she wants by talking about the options at age 65.

Jed tells me more about the options that I would have at age 65. I remark that I would use the options for purposes of retirement when I reached age 65. I am waiting for the chance to find out what amount of coverage I can get without paying more in monthly premiums.

26-30 Minutes

I know the sale is made when she said she would use her options for retirement purposes. But, I still don't know what amount. I believe about \$6,000 is needed. I decide on this as a goal. Kathy mentions disability income

I try to explain to Jed why I thought it was important to look at the disability income protection. I do this primarily because he has worked so hard at explaining both plans to me.

decided that the policy she was carrying was not appropriate to her needs. Jed called to set up an appointment in order to discuss changing her insurance coverage. Kathy had never met Jed before this time.

The Process Model

The verbal and graphical process models shown in Table 1 and Figure 1 reflect the prior knowledge of the buyer possessed by the salesperson. He had reviewed the buyer's insurance coverage and believed he had uncovered an unrecognized need. Thus, the salesperson had formulated a well-defined goal to convert the buyer's mortgage policy into a \$10,000 whole life policy. Kathy, on the other hand, had no prior knowledge of the salesperson and had no well defined goals or strategies. Vague goals of "finding out more about the salesperson" and "seeking information" were cited to the researcher.

Jed's strategy for accomplishing his goal was also explicit. His strategy required that he first get an appointment with the buyer; then establish rapport with his potential client; then show the buyer that the present policy carried was not needed; and finally convince the buyer that a whole policy would better meet her insurance needs. References to familiar others were relied on to establish rapport. Thus, for the salesperson, decision nodes 4, 6, 8 and 10 in Figure 1 were critical points in attempting to achieve his original goal.

Kathy attempted to accomplish her goal of learning about the salesperson during the first five minutes of the sales encounter. Questions directed to the salesperson for this purpose centered on (1) the salesperson's place of residence, (2) his place of education, and (3) persons known to both parties. This initial evaluation of the salesperson is reflected in decision node 3 in Figure 2. Such evaluations are consistent with social exchange theory.

The models in Figures 1 and 2 reflect the impact that actions by one party had on the other person's decision network. For example, a decision path compatible with the salesperson's original goal and strategy would have been: 1→2→4→5→6→7→8→9→10→22. The buyer, however, raised questions that required the salesperson to revise his original goal and strategy. Specifically, Kathy wanted to know (1) if the whole life policy would be compatible with other endowment policies (Figure 2, node 12), and (2) whether she should consider purchasing a disability income policy (Figure 2, node 13). The salesperson also reported detecting nonverbal cues as to Kathy's price expectations. This feedback represented a new array of cues to be faced by the salesperson, with the result that Jed's actual decision path was: 1→2→4→5→6→7→8→9→10→11→13→14→15→16→17→19→20→22. That is, the salesperson had to be flexible in changing his sales approach to include the selling of a disability income policy or a lesser valued whole life policy. Professional advice provided by Stan Trotter, the agency manager, to be flexible in the sales encounter was applicable with this buyer.

The salesperson, on the other hand, apparently exerted influence on what the buyer evaluated in her decision process. The salesperson initiated evaluations of whether the mortgage insurance was needed and whether a whole life policy should be considered as an alternative. The salesperson provided information on cash value and paid-up insurance features of the whole life policy that became part of the buyer's decision net. Finally, the buyer's purchase decision was made immediately following information on social security benefits that was provided by the salesperson (Table 1). Thus, decision nodes 5, 7, 9, 11, and 18 in the buyer's decision network may be attributed to the salesperson's influence.

Extensive problem solving behavior seems to characterize Kathy's decision process for insurance. She did not have

a well-formed set of choice criteria and had low familiarity and high search needs when considering an insurance purchase. She admitted to being confused with information provided by Jed on policy options. High search needs led Kathy to seek information from her brother-in-law and a friend. Both of these persons were contacted by the researcher, but no additional insight into their influence was gained. The brother-in-law did not advise against the whole life plan, but recommended that she explore the merits of a disability income policy first. The friend that Kathy contacted thought that she was worried that the insurance agent might be trying to sell her something unnecessary. He provided assurance that she need not worry. The basis on which this assurance was given was the reputation of the cooperating company.

In Kathy's decision net, nodes 18 and 20 were important turning points. As soon as Kathy discovered that she would be able to receive social security benefits if disabled, the purchase decision was made contingent upon her being able to keep monthly premiums below \$20. The salesperson did not detect her willingness to purchase and continued the sales presentation for approximately 15 minutes before attempting to close. Jed accurately perceived Kathy's concern for the price, however, and modified his own expectations to reflect this. Kathy admitted to the researcher that she did not want to tell the salesperson that she wanted to limit her premium to \$20 or less. Thus, she continued interaction without informing Jed of her purchase decision. This behavior suggests that the sale may have been lost if not for the correct reading of nonverbal cues and the willingness to be flexible by the salesperson. The reasons for Kathy's willingness to reveal her concern for the premium costs to the researcher and not to the salesperson were not determined.

Attempts by Jed at the end of the sales encounter to develop the long term relationship between customer and agent suggest that developmental selling may characterize the life and health insurance selling process. Jed's efforts were compatible with similar observations by the researcher of other salespersons.

Buyer-Seller Compatibility

Kathy was asked several questions concerning characteristics of the salesperson and the sales approach used in the retrospective analysis. Kathy described Jed as a person much like herself in respect to personality and values. When asked if he was similar in terms of social class and economic status, Kathy stated that she had not considered these factors. Kathy based her evaluations of Jed on information solicited as to where he lived and where he had gone to school.

Jed was considered by Kathy to be knowledgeable about insurance and a person to be trusted. His sales presentation was described as being unstructured and nonpressuring. Kathy particularly liked the concern she believed Jed expressed for her personally. She summarized her evaluation of Jed this way:

I do not think he was really out to sell the policy. I might have misinterpreted it, but he made me feel that the policy itself was for my betterment.

Process Models for Buying and Selling Centers B and C

Decision process models for two additional buying centers and two additional selling centers were constructed following one to two hours of retrospective analysis with each center. These models are not described here because of the complexity of the models and space requirements. Buyer center B consisted of two buyers,

Charles and Pam Trout. Unlike the case of Kathy Berryman, the Trouts met with the salesperson only one time and decided to purchase. Selling center B consisted of one salesperson. Excerpts of the interaction between the centers are reported elsewhere (Woodside, et al., 1977).

Selling center C consisted of two salespersons. Two buyers, Seaborn and Barbara Sellers, made up the buying center. A single sales encounter between these two centers took place in the Sellers home, the buying center consulted no outside sources before making a decision to purchase. Details of the process models for centers. B and C may be found elsewhere (Taylor, 1977).

The process models for centers B and C generally support the pattern of communication and problem solving described for center A. The results provide additional support to the Wilson and Bambic (1977) model of a multiple stage sequence of decision process in buyer-seller interactions. Wilson's model was formulated from the tenets of Homan's (1961) social exchange theory. Thus, the finding that the selling process proceeded through stages as posited by Wilson supports the view of life insurance buying-selling as social exchange. This is also supported by the observation that salespersons and buyers spent a relatively large amount of time engaged in social commentary that was unrelated to the product per se. Salespersons typically spent a relatively large part of the sales encounter in attempts to develop rapport and to establish a continuing relationship with the buyer. Similarly, buyers spent much effort attempting to learn more about the salesperson. In the retrospective analyses, the buyers emphasized the importance of salesperson characteristics on their willingness to purchase from a particular salesperson. One salesperson, on the other hand, informed the researchers that he did not try to sell to professional people because he never felt comfortable with professionals. To this salesperson, the social costs of interacting with businessmen, lawyers and other professionals apparently outweighed potential rewards of making a sale. Thus, the study findings reflect the importance of compatibility between buyer and seller such that the rewards of exchange are high relative to the social costs of exchange. Compatibility between buyer and seller was suggested by Sheth (1976) in his conceptualization of a successful buyer-seller interaction. Thus, the study also provides some support to Sheth's deductive model.

Implications

The feasibility of using recorded conversations in which participants know that they are being recorded was demonstrated in the study. Conversations with buyers following the sales encounter indicated that the researcher and recorder's presence was usually forgotten after a few minutes. Future research of this type, however, should attempt to supplement the audio recordings with video recordings. By doing so, measures of nonverbal communication behavior can be devised and results correlated with the buyer's verbal communication behavior. Results so obtained might lead to a richer understanding of the content of interaction between buyers and sellers.

This study conducted would be classed by Holbrook and O'Shaughnessy (1976) as interaction synchronization research; that is, the focus was on communications variables and a two-way perspective was taken. This is an improvement over past personal selling research that either focused on static characteristics of the buyer or seller or adopted a one-way perspective. Nevertheless, important independent variables in the study were not under the researcher's control. Future research of this type should attempt to manipulate the frequencies of using the bargaining and selling variables in order to ascertain the impact of their use on purchase. Sales-

persons might, for example, be provided with questionnaires that involve high and low levels of question usage. The different treatments (high versus low levels of asking questions) could be assigned to different customers in an attempt to evaluate their influence on sales success. Similar manipulations could be used for the other independent variables in the study.

A large number of salespersons and prospects should be used in future studies of this type. By using a larger number of buyers, more sophisticated statistical analyses can be conducted. Multiple regression could be used, for example, to determine the relative influences of the selling variables on purchase. Additionally, interaction effects between the independent variables could be studied. By using a larger number of salespersons, individual differences between salespersons can be examined; thus, leading to a greater confidence in generalizing study results.

REFERENCES

- Marcus Alexis, George H. Haines, and Leonard Simon, "Consumer Information Processing: The Case of Women's Clothing," Proceedings of the Fall Educators' Conference (Chicago: American Marketing Association, 1968), 197-205.
- Peter J. Bambic, "An Interpersonal Influence Study of Source Acceptance in the Industrial Buyer-Seller Exchange Process: An Experimental Approach," unpublished dissertation, The Pennsylvania State University, 1977.
- James R. Bettman, "Toward a Statistics for Consumer Decision Net Models," Journal of Consumer Research, 1 (February 1974), 71-80.
- Paul Busch and David T. Wilson, "An Experimental Analysis of a Salesman's Expert and Referent Bases of Social Power in the Buyer-Seller Dyad," Journal of Marketing Research, 13 13 (February 1976), 3-11.
- Noel Capon, Morris B. Holbrook, and James Hulbert, "The Selling Process: A Review of Research," in Arch G. Woodside, Jagdish N. Sheth, and Peter D. Bennett, eds., Consumer and Industrial Buying Behavior (New York: Elsevier North-Holland, 1977), 323-331.
- Richard M. Cyert, Herbert A. Simon, and Donald B. Trow, "Observation of a Business Decision," Journal of Business, 29 (April 1956), 237-248.
- R.M. Cyert and J.G. March, A Behavioral Theory of the Firm (Englewood Cliffs, New Jersey: Prentice-Hall, 1963).
- Franklin B. Evans, "Selling as a Dyadic Relationship: A New Approach," American Behavioral Scientist, 6 (May 1963), 76-79.
- M.S. Gadel, "Concentration by Salesman on Congenial Prospects," Journal of Marketing, 28 (April 1964), 64-66.
- Morris B. Holbrook and John O'Shaughnessy, "Influence Processes in Interpersonal Persuasion," in Advances in Consumer Research, Vol. III (Cincinnati: Association for Consumer Research, 1976), 364-369.
- George C. Homans, Social Behavior: Its Elementary Forms (New York: Harcourt Brace Jovanovich, 1961).
- John A. Howard and William M. Morgenroth, "Information Processing Models of Executive Decisions," Management Science, 14 (March 1969), 416-428.

John A. Howard, James Hulbert, and John U. Farley, "Organizational Analysis and Information-Systems Designs: A Decision Process Perspective," Journal of Business Research, 3 (April 1975), 133-148.

William M. Morgenroth, "A Method for Understanding Price Determinants," Journal of Marketing Research, 1 (August 1964), 17-26.

Allen Newell and Herbert A. Simon, Human Problem Solving (Englewood Cliffs, New Jersey: Prentice-Hall, 1972).

Richard W. Olshavsky, "Customer-Salesman Interaction in Appliance Retailing," Journal of Marketing Research, 10 (May 1973), 208-212.

Richard W. Olshavsky, "Consumer Decision Making in Naturalistic Settings: Salesman-Prospect Interaction," in Advances in Consumer Research, Vol. III (Cincinnati: Association for Consumer Research, 1976), 364-369.

Allen L. Pennington, "Customer-Salesman Bargaining Behavior in Retail Transactions," Journal of Marketing Research, 5 (August 1968), 255-262.

Jagdish N. Sheth, "Buyer-Seller Interaction: A Conceptual Framework," in Advances in Consumer Research, Vol. III (Cincinnati: Association for Consumer Research, 1976), 382-386.

James L. Taylor, "Exchange Behavior Among Life Insurance Selling and Buying Centers in Naturalistic Settings," unpublished dissertation, University of South Carolina, 1977.

W.T. Tucker, Foundations for a Theory of Consumer Behavior (New York: Holt, Rinehart, and Winston, 1967).

Ronald P. Willett and Allen Pennington, "Customer and Salesman: The Anatomy of Choice in a Retail Setting," Proceedings of the Fall Educators' Conference (Chicago: American Marketing Association, 1966), 598-616.

David T. Wilson and Peter J. Bambic, "Conceptual Models of Buyer/Seller Interaction," paper presented at the 85th Annual Convention, American Psychological Association, San Francisco, 1977.

Arch G. Woodside and Robert A. Fleck, Jr., "Consumer Choice Processes-An Intensive Study of Two Cases," paper presented at the 85th Annual Convention, American Psychological Association, San Francisco, 1977.

Arch G. Woodside, James L. Taylor, S. Travis Pritchett, and William M. Morgenroth, "Transactions Among Buying and Selling Centers," in Proceedings of the Fall Educators' Conference (Chicago: American Marketing Association, 1977), 59-64.

NARROWING THE CONCEPTUAL-EMPIRICAL GAP IN ORGANIZATIONAL BUYING BEHAVIOR

Robert E. Spekman, University of Maryland
Bobby J. Calder, Northwestern University

Abstract

Marketing theorists depict the organizational buying process as comprised of several organizational members, each participating, to a greater or lesser extent, in the purchasing decision. At the same time, however, there is a noticeable lack of empirical research directed at the collective level of analysis with the goal of capturing the interpersonal dynamics of the procurement process. This paper attempts to narrow the gap between conceptualizing the process of joint decision-making and empirically investigating aspects of the collectivity responsible for purchasing related decisions. Specifically, a macro-sociological and a structural role approach are discussed as two feasible methodologies.

Introduction

In an attempt to develop an historical perspective to the study of organizational buying behavior from the aggregated, or collective, level of analysis an early study by Strauss (1962) will serve as a point of departure. Strauss, capturing the bureaucratic infighting between purchasing and other functional areas, portrayed organizational procurement as a political process in which each participant attempts to influence the purchasing decision. Interestingly, however, a great deal of the subsequent research (e.g., Leavitt, 1967; Lehmann and O'Shaughnessy, 1974; Wilson, 1971; Wind, 1970) focused solely on a key organizational member, such as the purchasing agent, and examined the extent to which various factors affected this individual's decision to select a particular vendor. While the managerial implications gleaned from an investigation of individual preferences and differences may further the development of micro-segmentation strategies (Wind and Cardozo, 1974), it is probably safe to conclude that an individual level of analysis is not very useful in capturing the dynamic interactions among those organizational members who share in the procurement process.

Beginning with descriptive studies by Robinson, Faris and Wind (1967), other researchers (Feldman and Cardozo, 1966; Grønhaug, 1976; Patchen, 1974; Sheth, 1973; Weigand, 1968) have built upon Strauss' representation of organizational buying as a joint decision making process. Webster and Wind (1972), in their influential analysis of organizational buying behavior, developed the concept of the buying center and thereby attempted to raise the level of analysis to include that collectivity responsible for purchasing related decisions. Embodying Strauss' conceptualization of the buying process, the buying center notion focuses on those formal and informal relationships which develop between the purchasing agent and those other organizational actors with whom he interacts during the purchasing process. It should be noted that unlike a formal structural subunit (i.e., the purchasing department or the marketing department), the buying center is a more nebulous construct, reaching across functional boundaries, whose composition can only be determined through empirical investigation.

In light of the realities of organizational decision-making (see, Cyert and March, 1963), a collective level of analysis would seem to more accurately reflect the dynamic interactions which depict the organizational

procurement process. Why, then, is there a paucity of empirical research focusing on a collective level of analysis? An explanation appears to lie in the lack of an appropriate methodology. The purpose of this paper is to narrow the gap between conceptualizing the process of joint decision-making and empirically analyzing aspects of the collectivity referred to as the buying center. To achieve this goal we will explicate two methods for conducting research which reveal information about how the activities of many individuals come together to affect purchasing decisions. One approach emerges from macro-sociology and is concerned with structural properties of the buying center. The second approach evolves from role theory and emphasizes the role structure of those individuals composing the buying center.

Structural Properties of the Buying Center

Background

Structural properties of the buying center are obtained by analyzing data about the relations of each buying center member to some or all of its members (Lazarsfeld and Menzel, 1969). With the advent of open systems thinking (Kast and Rosenzweig, 1972), some organizational theorists (Burns and Stalker, 1959; Duncan, 1972; Galbraith, 1973; Hall, 1963; Lawrence and Lorsch, 1967) began to view the organization (as well as its subunits) as dynamic, responsive entities in continual, and necessary, interaction with its environment. In short, the organization became an information processing entity and its structural configuration defined its ability to gather and process information. Breaking from the Weberian tradition, these contingency researchers posited that some structural configurations are more adept than others for dealing with the varied environmental factors with which the organization must contend. Moreover, a basic premise of this stream of research is that of all the environmental factors with which a firm (or one of its subunits, for that matter) must contend the most pervasive and all encompassing is environmental uncertainty. By adapting its structural configuration to match the level of uncertainty in its environment, a firm (or a subunit) can facilitate the gathering and processing of information crucial to its decision-making process.¹

Dimensions of Structure

Following a contingency related paradigm, the buying center is conceived of as a "decision unit" whose members are charged with making purchasing related decisions (Spekman, 1977a). In this fashion, the buying center exists as a communication network which does not necessarily derive its structural configuration from the formal organizational per se; but rather from the regularized patterning of behavior and communication flows which typify the industrial procurement process. In a very real sense, the buying center's structural configuration also serves to define its decision-making

¹It should be noted that while uncertainty has been portrayed as an independent variable here, other researchers have relied on technology (Mohr, 1971; Woodward, 1965), size (Child, 1972) and other environmental factors as antecedents to the structural dependent variables.

potential. Extant contingency research (Duncan, 1972; Hall, 1963) has, for the most part, concentrated on four dimensions of structure: 1) centralization--the degree to which authority and responsibility rests with a particular organization member; 2) rules and procedures--the extent to which activity in an organization is formally defined; 3) division of labor--the degree to which tasks are differentiated; and 4) participation in decision making--the extent to which organizational members are involved in decision-making. From an analysis of these structural measures one is able to generate a profile of the interactions which take place in the buying center during the decision process.

As stated earlier, some structural configurations are more conducive to information gathering and processing than other forms. Specifically, a highly bureaucratized structure restricts the flow of information to rigidly established channels of communication and is best suited for decision-making when 1) time is of the essence; 2) the decision is routine in nature; and 3) environmental demands are clear and uncomplicated (Katz and Kahn, 1966). Clearly, it can be seen that a highly bureaucratized structure--with its high degree of centralization, strict rules and procedures, clearly defined and distinct division of labor, and low levels of participation in decision-making--places severe limitations on the amounts of information available to the decision unit (i.e., the buying center).

As the decision-making requirements become more complex--perhaps as a result of greater environmental uncertainty--empirical evidence (Duncan, 1972; Lawrence and Lorsch, 1967) would suggest that a less bureaucratized structure would improve the information gathering and processing capabilities of the decision unit. For instance, a lower level of centralization would promote decision-making at the point of information gathering rather than restricting the flow of information to a more rigidly defined communication path. Further, the content of purchasing-related communication would consist of greater information passage and less instructions.

Similarly, a less rigid division of labor enables the decision unit to remain flexible enough to allow for a constant redefinition of tasks which may be necessitated by the demands of a highly complex, non-routine decision-making situation. Through increased lateral communication and a more flexible design, the buying center members can more quickly and more easily respond to changing demands and contingencies. Furthermore, flexible rules and procedures enhance effectual processing of novel and non-routine information by permitting greater individual discretion. Such structural adaptation can lead to innovative behavior on the part of the buying center members and, thus, allow them to adapt to contingencies and information requirements which may not have been foreseen when the rules and procedures were initially developed.

The last structural dimension concerns the degree of participation in decision-making. Greater participation is an outgrowth of the adaptation associated with the previous structural dimensions. It logically should follow that as the buying center members are faced with more difficult decisions greater dependency should be placed on informal, lateral channels of communication. This increased participation in decision-making should improve the decision-making potential of the buying center members by enlarging the information base available to them as well as by providing opportunities for more sources of information feedback (Argyris, 1973).

From the dimensions of structure outlined above one can then position a buying center along a bureaucratic/non-bureaucratic continuum. A particular buying center

profile would presumably connote a particular communication network, concentration of authority, workflow and so on among the buying center members. These, and other, organizational factors have been shown to impact upon aspects of the buying process (Grønhaug, 1976; Webster and Wind, 1972). While superior to Webster and Wind's descriptive, functional role approach, a possible flaw in these structural measures, as presently operationalized, is that these constructs may not accurately reflect the informal interactions among the buying center members. For example, it is possible to infer that a less bureaucratized, more flexible buying center profile represents a more democratic, participative decision-making atmosphere with a high degree of shared purchasing responsibilities and a de-emphasis on impersonal behavior. Yet, such organizational level analyses do not readily permit the researcher to examine the precise nature of the various sociometric linkages which emerge during the procurement process.

It should also be mentioned that the aggregation procedure itself tends to mask any individual level behavior since the buying center profile reflects an average of the individual responses. Moreover, recent findings suggest that there is a relatively low degree of consensus among the various buying center members with respect to their perceptions of the buying center structure (Spekman, 1977b). Despite the inflated error variance associated with each buying center score resulting from the above instability (Cook and Campbell, 1976), the macro-sociological approach does furnish insights pertinent to the decision-making potential of the buying center members. Specifically, it has been shown that the various structural dimensions have important implications for the decision unit's ability to gather and process information. Nonetheless, what may be more relevant to the researcher is a methodology which focuses on an aggregated, or collective, level of analysis and, as importantly, also facilitates the examination of the individual as he interacts as a member of the particular collectivity under examination.

Structural Role Analysis

Role theory provides a conceptual framework for linking an individual to a collectivity and, as Hickson (1966) suggests, it is a useful vehicle for bridging the organizational and individual level of analysis. Hall (1976) states that role theory is a helpful device for understanding the interplay among persons, interpersonal and organizational factors as they impinge upon the behavior of organizational members. Structural role analysis, a subset of role theory, concerns itself primarily with the structure of role relationship. In comparison to the macro-sociological approach, structural role analysis is a more micro-level orientation in that it calls for an intensive analysis of the individual decision unit or organization.

Elements of Role Structure

Basic to the concept of role structure is that any collectivity is linked by various tasks which must be accomplished. The particular workflow which emerges establishes specific relationships among the members of the collectivity. In the context of the buying center, for instance, there are generalized decision stages--ranging from identification of a buying need to performance feedback and evaluation--which must be performed during the procurement process. While the exact nature and sequence of the decision states (i.e., BUY-PHASES) may vary depending upon the type of purchase (i.e., straight rebuy, modified rebuy or novel purchase) there is usually a clearly established workflow which can be discerned.

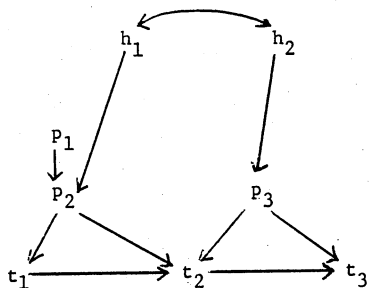
A second element, the concept of positions, links a particular person to a task element. If, for example, one of the decision stages in the procurement process is the setting of specifications, we may find that the individual performing this task is called a design engineer--regardless of who the individual is, the job is always done by a design engineer. The notion of positions is quite consistent with the impersonal, Weberian notion of organization. In this manner, the set of linkages illustrated by the various positions within the buying center serves to 1) establish the formal relationships among buying center members and 2) represent the functional role (i.e., design engineer, buyer, etc.) assigned to a particular task. Simply, the concept of position defines one's position on an organizational chart and specifies the nature of one's formal interactions with other members of the organization or organizational subunit (in this case, the buying center).

Since the formal structure does not, in many cases, depict all the interactions within an organization, a third set of relationships emerges. The third set of elements concerns persons who are individuals initially interrelated by their assignment to a particular position and who, over time, can develop informal relationships which need not be specified by the formal hierarchy of the organization. It can be seen that this set of persons serves to complete the communications network (Oeser and Harary, 1962 and 1964) not fully accounted for by the more formal delineation of positions or tasks.

Digraphs, roughly defined as patterns of relationships among pairs of abstract elements (Harary, Norman and Cartwright, 1965), are used to illustrate and describe how any set of tasks, positions, and persons are conceived to be logically interrelated. While Figure 1 maps out a completed structural role system, it is possible to isolate a series of graphs where a P-graph depicts the formal relationships among buying center members, a T-graph outlines the sequence of purchasing related tasks, and an H-graph delineates the informal relationships among the buying center members. From Figure 1 one can discern 1) a sequence of buying activity (beginning with t_1 and ending with t_3); 2) a superior-subordinate relationship between a buyer (p_2) and a purchasing manager (p_1); and 3) an informal relationship between individuals acting as a buyer (h_1) and a design engineer (h_2) both of whom share the task of setting a particular product's specification (t_2). While no doubt a simplified structural role relationship, one can easily envision quite an elaborate set of digraphs encompassing many individuals, occupying several organizational positions, and engaging in a detailed sequence of procurement activities.

Figure 1

A Simplified Digraph Depicting a Structural Role System in a Buying Center



From a highly schematic form of digraph one is able to more deeply delve into the interpersonal dynamics of the buying center. Not only can one more easily discern the sequence of purchasing related events and the various positions sharing these activities, a researcher is able to trace informal power relationships and other sociometric linkages which cannot be examined directly from a more macro-perspective. A structural role approach offers a basis for aggregating across organizations while still preserving an individual, micro-level perspective.

It should be noted that the use of structural role analysis in an organizational buying context is presently not well developed (Calder, 1976) and, therefore, it is not our intention to propose it as a panacea for existing methodological weaknesses. Rather, our purpose has been to portray structural role analysis as a useful conceptual framework for capturing the micro-level relationships which lie at the core of the buying center construct. Utilizing a structural role perspective, we have, in fact, just finished collecting data pertinent to the decision-making properties of organizational buying centers. Yet, we sought to capture the actual micro-level relationships among buying center members through a multidimensional scaling technique. The point to be made is that structural role analysis is not a method in and of itself; it merely provides a guide by which to conduct collective level analyses.

Conclusions and Implications

The issue of whether the individual or the collectivity is the proper level of analysis for conducting inquiry into organizational buying behavior is one with which future researchers must come to grips. While it is quite acceptable to empirically deal at both levels, it is felt that for much of the information germane to the joint decision-making properties of the collectivity responsible for purchasing related decisions (i.e., the buying center) the aggregated level of analysis is appropriate. For the present, the major stumbling block appears to be the lack of a suitable methodology.

The objective of this paper has been to stress the need for narrowing the gap between the conceptualization of organizational buying as a joint decision-making process and the method by which a researcher can empirically investigate that process. While still in the exploratory stages both the macro-sociological approach and the structural role approach appear to be viable methods for capturing aspects of the joint decision-making behavior of the buying center members. Moreover, both approaches permit the researcher to examine relationships among variables both endogenous and exogenous to the system thereby surmounting many of the pitfalls of investigating the individual members of the buying center as isolated entities.

The macro-sociological approach, however, primarily allows the researcher to infer particular interpersonal relationships from the structural configuration of the collectivity under scrutiny while the structural role approach can result in a more revealing picture of the interpersonal dynamics of the buying process. That is, much more verbal information can be gleaned from a graphic representation of the various structural role elements than is available from a macro-sociological investigation. It may be possible, for example, to infer from a rather flexible structural configuration that a purchasing agent is involved in several stages of the purchasing decision as a result of the more participative style of decision-making. Yet, a structural role analysis of the same buying center may disclose that this same purchasing agent influences particular decision stages (perhaps, the setting of

specifications) through his informal, social relationships with a particular design engineer. Since competitive suppliers may only meet with the purchasing agent such information has important implications for the future conduct of their marketing efforts. Clearly, one can envision the strategy implications which can be distilled from a thorough analysis of organizational buying behavior from the collective level of analysis. We have presented two feasible directions for empirically examining aspects of organizational buying behavior from the collective level of analysis.

References

- Chris Argyris, Intervention Theory and Method (Reading, Ma.: Addison-Wesley, 1973).
- Tom Burns and George Stalker, The Management of Innovation. (London: Tavistock Publishing Co., 1959).
- Bobby J. Calder, "Structural Role Analysis of Organizational Buying," presented at Symposium on Consumer and Industrial Buying Behavior, University of South Carolina, 1976.
- John Child, "Organizational Structure, Environment and Performance," Sociology, 6 (1972), 1-22.
- Thomas Cook and Donald T. Campbell, "The Design and Conduct of Quasi-Experiments and True Experiments in Field Settings," in M. D. Dunnette (ed.). The Handbook of Industrial and Organizational Psychology (Chicago: Rand McNally, 1976).
- Richard Cyert and James March, The Behavioral Theory of the Firm (Englewood Cliffs: Prentice-Hall, 1963).
- Robert B. Duncan, "Characteristics of Organizational Environments and Perceived Environmental Uncertainty," Administrative Science Quarterly, 17 (1972), 313-27.
- William Feldman and Richard Cardozo, "The Industrial Revolution and Models of Buyer Behavior," Journal of Purchasing, 5 (November, 1966), 77-88.
- Jay Galbraith, Designing Complex Organizations (Reading, Ma.: Addison-Wesley, 1973).
- Kjell Grønhaug, "Exploring Environmental Influences in Organizational Buying," Journal of Marketing Research, 13 (August, 1976), 225-29.
- Richard Hall, "The Concept of Bureaucracy," American Journal of Sociology, 69 (1963), 32-42.
- _____, Organizations: Structure and Process (Englewood Cliffs: Prentice-Hall, 1976).
- Frank Harary, R. Norman, and D. Cartwright, Structural Models: An Introduction to the Theory of Directed Graphs (New York: Wiley, 1965).
- David Hickson, "A Convergence in Organizational Theory," Administrative Science Quarterly, 11 (September, 1966), 225-37.
- F. Kast and L. Rosenzweig, "General Systems Theory: Applications for Organizational Management," Journal of the Academy of Management, 15 (December, 1972), 447-468.
- Daniel Katz and Robert Kahn, The Social Psychology of Organizations (New York: Wiley, 1966).
- Paul Lawrence and Jay Lorsch, Organization and Environment (Homewood, Il.: Richard D. Irwin, 1967).
- Paul Lazarsfeld and H. Menzel, "On Relations Between Individual and Collective Properties," in A. Etzioni (ed.). A Sociological Reader of Complex Organizations (New York: Holt, Rinehart & Winston, 1969).
- Theodore Leavitt, Industrial Purchasing Behavior (Boston: Graduate School of Business Administration, Harvard University, 1965).
- Donald Lehmann and John O'Shaughnessy, "Difference in Attribute Importance for Different Industrial Purchases," Journal of Marketing, 38 (April, 1974), 36-42.
- Lawrence Mohr, "Organizational Technology and Organizational Structures," Administrative Science Quarterly, 14 (December, 1971), 444-59.
- O. Oeser and F. Harary, "A Mathematical Model for Structural Role Theory, I," Human Relations, 15 (1962), 89-109.
- _____, "A Mathematical Model for Structural Role Theory, II," Human Relations, 17 (1964), 3-17.
- Martin Patchen, "The Locus and Basis of Influence on Organizational Decisions," Organizational Behavior and Human Performance, 11 (1974), 195-221.
- Patrick Robinson, Charles Faris, and Yoram Wind, Industrial Buying and Creative Marketing (Boston: Allyn & Bacon, 1967).
- J. Sheth, "A Model of Industrial Buyer Behavior," Journal of Marketing, 37 (1973), 50-56.
- Robert E. Spekman, "A Contingency Approach to Power Relationships Within the Organizational Buying Task Group," Unpublished Ph.D. Dissertation, Northwestern University, 1977a.
- _____, "An Examination of Informal Decision-Making Units: Some Methodological Cautions for Future Research," Proceedings of the Academy of Management, (1977b), 218-23.
- George Strauss, "Tactics of Lateral Relationships," Administrative Science Quarterly, 7 (September, 1962), 161-86.
- Frederick Webster and Yoram Wind, Organizational Buying Behavior (Englewood Cliffs: Prentice-Hall, 1972).
- Richard Weigand, "Why Studying the Purchasing Agent Is Not Enough," Journal of Marketing, 32 (1968), 41-45.
- David Wilson, "Industrial Buyers' Decision-Making Styles," Journal of Marketing Research, 8 (1971) 433-36.
- Yoram Wind, "Industrial Source Loyalty," Journal of Marketing Research, 7 (1970), 450-57.
- _____, and Richard Cardozo, "Industrial Marketing Segmentation," Industrial Marketing Management, 3 (1974), 153-66.
- Joan Woodward, Industrial Organization: Theory and Practice (London: Oxford University Press, 1965).

ON THE INTERFACE BETWEEN ORGANIZATIONAL AND
CONSUMER BUYING BEHAVIOR

Yoram Wind, University of Pennsylvania

Abstract

The relevance of organizational buying behavior concepts and methods to the study of consumer behavior is examined. The paper suggests a number of implications for the study of consumer behavior and calls for the legitimization of the study of organizational buying behavior within the boundaries of the consumer behavior discipline.

Introduction

The basic premise of this paper is that our concepts of and approaches to the study of consumer behavior can benefit from a close examination of the organizational buying behavior literature.¹ Many of the consumer behavior models recognize the social context and multiperson nature of most purchase and consumption behavior. Yet, to date, these models and most of the consumer behavior concepts and studies have focused on the individual as the unit of analysis.

The lopsided emphasis on the individual and the tendency to ignore the multiperson issue can be attributed to three major reasons:

1. Conceptual difficulties involving the development of multiperson variables (as distinct from individual variables) and hypotheses.
2. Methodological difficulties involving the analysis of multiperson data.
3. Operational difficulties involving added time and monetary costs for collection and analysis of multiperson data.

Organizational buying behavior, on the other hand, did not have the luxury of focusing on the individual. Even the early conceptual and empirical organizational buying behavior studies recognized the fact that meaningful studies cannot focus only on the purchasing agent but rather should be concerned also with other organizational members (e.g., users) involved with the purchase and usage of the given products or services.

In this respect then, the usually neglected area of organizational buying behavior could conceivably provide concepts and methods relevant to the study of multiperson consumer behavior. The objective of this paper is to briefly outline some of these possible contributions and suggest that advances in our understanding of consumer behavior can be gained not only from research on consumer behavior but also indirectly from research on organizational buying behavior.

Conceptual Contributions

The conceptual contributions of organizational buying behavior to the study of consumer behavior include the concepts of:

¹The converse (i.e., the contribution of consumer behavior models, concepts, and findings to the study of organizational buying behavior) is also true. Yet, given the orientation of this audience, the paper focuses on the contributions of the organizational buying behavior area to the study of consumer behavior.

- . the buying center
- . decision stages and roles in the buying center
- . group choice models
- . buying situations
- . a building block modeling approach
- . buying center tasks and objectives
- . consumers' organizational climate and structure
- . consumers' organizational technology
- . consumers' organizational (vs. individual) variables

The Buying Center

Consumer products and services are purchased not only by individuals, but quite frequently by a number of individuals (not all of whom are members of the same household) or by an individual who buys for the consumption of someone else. Even when a person buys a product or service for his or her own consumption, a number of aspects of the purchase decision are often influenced by some relevant others. In all of these cases a sole focus on the individual could be quite misleading.

Recognizing the multiperson nature of most consumer behavior, greater emphasis has been placed in recent years on the study of the family and, in particular, the husband-wife dyad. Yet, the husband-wife dyad is not always the relevant unit of analysis (consider, for example, purchases involving products for children and the cases in which consumers are not living within the boundaries of the traditional family). It would be much more desirable, therefore, to accept the concept of the buying center, which was developed in the context of organizational buying behavior (Wind, 1967). The concept calls for the identification, in each purchase situation, of the relevant individuals involved in the buying decision process (whether they are members of the same household or not) and using them (the members of the buying center) as the relevant unit of analysis.

Conceptually, the buying center is the appropriate unit of analysis in consumer behavior. Yet, implementing the concept requires the resolution of a number of conceptual and methodological issues (Wind, 1977). A number of these issues are currently studied in the context of organizational buying behavior, and could provide better insight into consumer behavior.

Decision Stages and Roles in the Buying Center

Consumer behavior involves a large number of decisions² ranging from the identification of a need, through various stages of generation and evaluation of alternatives,

²A large number of models have been proposed in both the consumer and organizational behavior literature to describe the various buying decisions and their sequence. Since there does not seem to be any consensus as to the "true" sequence of purchase decisions, the readers may select the model of their choice replacing the rows in Exhibit 1 with the decision stages of their own model.

to the selection of a product, specific features of a given brand, and time and place of purchase. Since these decisions are typically made by a number of members in the buying center, it is useful to construct a matrix of decision stages by roles as illustrated in Exhibit 1 (which is an adaptation of a matrix developed in an industrial buying study). The specific decision stages and roles would most likely change across buying situations. The same individual could occupy a number of roles (e.g., buyer and user), and a number of individuals could occupy the same role (e.g., there may be a number of users).

EXHIBIT 1
Decision Stages by Roles³

DECISION STAGES	ROLES ⁴				
	Buyer	User	Influ- encer	De- cider	Gate- keeper
Identification of needs					
Establish specifications					
Search for alternatives					
Establish contact					
Set purchase and usage criteria					
Evaluate alternative buying actions					
Determine budget availability					
Evaluate specific alternatives					
Negotiate with suppliers					
Buy					
Use					
Conduct post-purchase evaluation					

Developing the relevant decision stages-by-roles matrix and establishing common patterns across various buying situations would help our understanding of consumer behavior and place the work on the individual consumer's decision processes in the proper perspective.

Furthermore, in many of these decision stages the person occupying a given role engages in a negotiative process with suppliers and other external personnel. In these cases the unit of analysis can be changed to the group of boundary persons (e.g., buyer and salesman).

Choice Models

The consumer behavior literature has been predominantly occupied with individual choice models. Yet, many consumer behavior situations involve a group choice⁵ which cannot be explained adequately by individual choice models.

The organizational buying behavior literature and its supporting disciplines (e.g., organizational behavior

³For alternate formulation of decision stages-by-roles matrix see, for example, Webster and Wind (1972).

⁴Roles can also be defined as husband, wife, children, friends, etc.

⁵The other case of individual choice subject to the influence of relevant others (Wind, 1976) can be viewed as being somewhat between the two extremes of the "pure" individual choice and the group choice situation.

and political science) have offered a number of group choice models. These models cover both formal "voting" type choice models and informal interaction type models. The formal group choice models vary with respect to the specific decision rules employed (e.g., simple majority), the way these rules are determined, and the weight given each group member. The informal interaction choice models vary with respect to the processes of interaction, influence, persuasion, bargaining, and conflict resolution procedures employed.⁶

Buying Situations

One of the early empirical organizational buying behavior studies (Robinson and Faris, 1967) distinguished between three major buying situations—the new task, modified rebuy, and straight rebuy situations. These situations vary with respect to the decision process and information inputs they require as well as with respect to the number of new alternatives considered and uncertainty involved.

Understanding the buying situation is as critical in the study of consumer behavior as it is in the study of organizational buying behavior. Yet, with the notable exception of Howard's (1963 and 1977) classification of consumer behavior into a similar trichotomy of routines, limited and extensive problem solving, and the theoretical and empirical work of him and his colleagues, little explicit attention has been given in the consumer research area to the importance of buying situations in the explanation of consumer research.⁷

A Building Block Modeling Approach

Whereas most consumer behavior models have attempted to specify the exact network of buying decisions, their interrelationships and determinants (e.g., Howard and Sheth (1969)), the organizational buying behavior model proposed by Webster and Wind (1972) is based on a building block approach, i.e., the model focuses on the identification of the major sets (blocks) of variables that could affect the organizational buying decisions.⁸ Recognizing that no deterministic model can explain and predict accurately all organizational buying decisions, the model provides a framework for the identification of variables and specification of hypotheses (reflecting the current state of knowledge) on the relationship among the key variables. As such, it serves only as a guideline for the selection of variables for specific empirical studies on organizational buying behavior with no attempt to specify general cause and effect type relationships.

Given the disappointing results (in terms of predictive efficacy) of most of the detailed consumer behavior models, it might be useful to lower our level of expectations from these models and reformulate them as building block models of consumer behavior. Alternatively, one can even take the organizational buying behavior block

⁶For a detailed discussion of a number of these group choice models see, for example, Patanik (1971), Fishburn (1973), and Guetzkow and Collins (1964). A comparison of the predictive ability of a number of these models in the context of both an organizational and family choice setting is being conducted by Buss (1977).

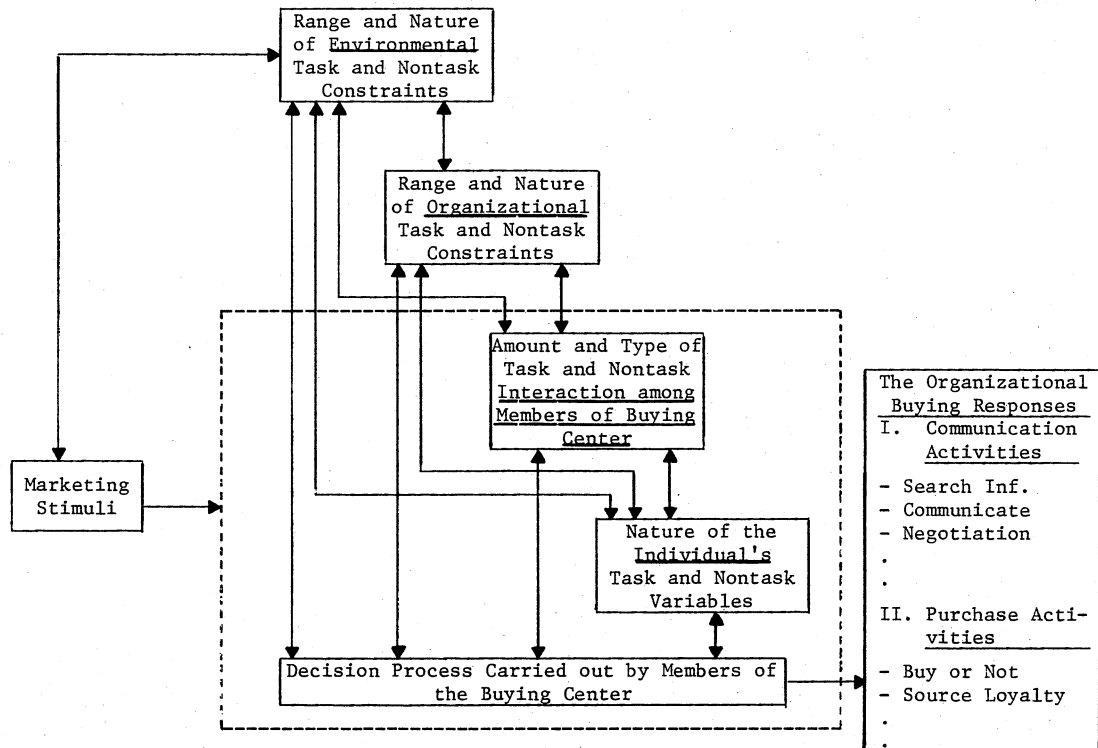
⁷Buying situations have also been approached from a different point of view by Belk (1975), who focused on the various dimensions of the consumer situational characteristics; namely, the physical and social surroundings, the temporal perspective, the task definition, and the antecedent states.

⁸For other modeling approaches to organizational buying behavior, see Sheth (1973), Hillier (1975), and Choffray and Lilien (1977).

model presented in Exhibit 2 (Webster and Wind, 1972) and use it in some modified form as a guideline for consumer behavior studies.

tural characteristics which have been ignored by the more conventional approaches to the study of consumer

EXHIBIT 2
An Organizational Buying Behavior Block Model



Buying Center Tasks and Objectives

The consumer behavior literature has given considerable attention to consumers' objectives as defined by their needs or benefits sought. Yet, little attention has been given to the identification of consumers' purchase related tasks, their relationship to the consumers' hierarchy of objectives, the ways consumers attempt to resolve conflicting objectives, and the evaluation of the purchase tasks (and functions).

It is in this context that the organizational buying behavior and procurement literature can be of help since one of their major concerns is with these issues (e.g., value analysis, economic order quantity procedures, etc.) and their impact on the buying process.

Consumers' Organizational⁹ Climate and Structure

The "organizational climate" literature distinguishes nine dimensions of organizational climate (Litwin and Stringer, 1968) which could also be used to describe the relevant consumer climate. These dimensions are: structure, responsibility, reward, risk, warmth, support, standards, conflict, and identity. The identification of these dimensions in the context of consumer behavior can add a rich new perspective to the more conventional consumer psychographic inventories.

Some of these dimensions are also related to the struc-

⁹The term consumers' organization is used to denote the group of people comprising the organization within which the consumers' purchase and consumption activities take place. The family, for example, can be considered such an organization.

behavior. Among the more relevant structural characteristics are:

The reward system. In the organizational buying context the behavior of individual buying center members was explained by their perceptions of the criteria used to distribute rewards (Wind, 1971). It seems, therefore, that in multiperson consumer buying centers the consumers' perceived rewards should be examined as one possible explanation (and predictor) of purchase behavior.

Division of responsibilities. The division of responsibilities within the consumer organization (the actual matrix of decisions by roles) and its stability (over time and situations) is one of the major factors which effect organizational buying behavior. It might be useful, therefore, to consider this variable (and its related concepts—the degree of decentralization and profit responsibilities of the buying function) in consumer studies as well.

Roles, role expectation, and role conflict. Organizational buying behavior and the sociological literature have suggested the importance of perceived roles, role expectations, and role conflicts in explaining the behavior of organizational members. The same concepts seem to be appropriate for the explanation of consumer purchase (and usage) behavior.

Authority structure. The buying authority and the extent of lateral vs. vertical involvement in the purchase process have long been recognized in organizational buying (for example, Strauss (1962)). These concepts have largely been ignored in consumer behavior despite the similarity between the authoritarian-democratic structure of the family and that of an organization. Simi-

larly, the purchase involvement of two parents (lateral relationship in egalitarian type family) differs from that of a parent and a child (vertical relationship).

Consumers' Organization Technology

It is well recognized in the organizational buying literature that technology has its impact both on the determination of which item to buy and on the nature of the buying process itself. It is reasonable to assume that technology is also a major factor in consumer behavior. Yet, it has been largely ignored in the study of consumer behavior.

In the consumer context, technology includes the house and equipment owned and used by the consumer, as well as the programs and procedures used to "manage" the household affairs in general and the buying function in particular.

The consumer's technology defines the technological constraints within which purchase and usage decisions and behavior take place. It is important to remember that the current house and equipment, as well as the technological orientation of the consumer, places significant constraints on the buying actions that can be considered. It might be useful, therefore, to examine consumer behavior conditional on the consumers' technological constraints.¹⁰ This is especially crucial for new products which either have to fit the consumers' current state of technology or change it (most likely at a high-er psychological and monetary cost).

Telephone and mail buying are examples of buying technology. Other technological developments should also be given attention. In particular, two-way cable TV and the recently introduced home computers which, at a price equivalent to a color TV set, have the potential of changing consumer buying technology, and affecting the buying decision process and outcomes.

Consumer Organizational (vs. Individual) Variables

The dependent variable in many consumer behavior studies has measured a response of the consumer organization (e.g., household purchase) and not of the individual decision maker. In contrast, the explanatory variables used in consumer behavior studies tend to be, with few exceptions (e.g., family income), individual in nature (e.g., education, sex, personality, attitudes, etc.).

Viewing the consumer as an organization suggests the addition of a new class of variables—namely, "organizational" type variables which measure the characteristics of the relevant consumer organization (apart from the characteristics of the individuals involved). The development of a set of consumer "organizational" measures requires a theoretical framework which encompasses the relevant dimensions of consumers' organizational behavior. In the absence of such a framework (which also does not exist for organizational buying situations) one may consider a number of variables¹¹ (and their possible consumer behavior descriptors) such as:

Domain consensus. (Levine and White, 1961)—the degree to which members of the consumer organization (e.g., household) agree on the goals, reference orientation, and tasks.

¹⁰This (technological) conditional analysis is similar in spirit to the one in which purchase and usage behavior is explained conditional on the consumers' existing inventory (assortment) or products (Wind, 1977c).

¹¹For a detailed discussion of organizational measures, see Price (1972).

Stability. (Caplow, 1964)—the length of time the consumer organization has been in its current structure (e.g., length of stay in a family life cycle stage, etc.).

Resource distribution. (Evan, 1966)—the amount and type of resources held by each member of the buying center compared with their needs.

Formalization. (Hage and Aiken, 1970)—the degree to which a consumer organization has rules and established guidelines for its purchase and consumption behavior.

Satisfaction. (Smith, Kendall, and Hulin, 1969)—the consumer's satisfaction with other members of the buying center (those occupying parallel roles to "co-workers" and "supervisors").

Efficiency. (Etzioni, 1964)—the amount of resources used by the consumer organization to achieve its intended purchase and consumption goals.

Intimacy. (Hemphill and Westie, 1950)—the degree to which members of the buying center are familiar with the preferences of the other members and know their probable reaction under widely different circumstances.

Participation. (Hemphill and Westie, 1950)—the degree to which members of the buying center apply time and effort to the purchase and consumption activities of the buying center.

In addition to the "new" consumer organizational type independent variables that the study of organizational buying behavior suggests for inclusion in consumer behavior models, there are a number of concepts suggested by the organizational buying behavior literature which should receive greater attention by consumer researchers. These concepts include:

The make, buy, or lease decision. Most consumer behavior studies focus on the buy decision ignoring the other two alternatives facing the consumer—leasing or making the product (or performing the service) by themselves.

The dyadic relationship between buyer and seller. Whereas most consumer studies focus on the consumer, a number of organizational buying studies have attempted to examine the nature of interactions between the buyer-seller, their negotiations, and bargaining processes. Given that in many consumer situations (e.g., in the retail environment) the consumer interacts directly with a seller, it might be useful to borrow from this area of the organizational buying behavior literature.

Multiple sourcing. Most organizations have policies against single source and attempt to maintain at any given time a number of alternative sources of supply. The consumer behavior literature has ignored this phenomenon and focused in many studies on the selection of a single brand or source. There is, however, some evidence to suggest that consumers do buy multiple brands from multiple sources. This was especially evident during the energy crisis when consumers tried to maintain "good customer relations" with more than a single gas station.

Methodological Contributions

Comparing the level of sophistication and innovativeness of consumer and organizational buying research suggests a clear dominance of consumer research.¹² Yet, despite

¹²For a discussion of some of the limitations of consumer research, see Jacoby (1976) and Wind (1977b).

the more advanced state of research on consumer behavior, there are a number of methodological contributions the organizational buying behavior literature can make to the advancement of consumer research. These contributions encompass areas which received considerably more attention in the study of organizational buying behavior and which could be applied to the study of consumer behavior. In particular these include:

- . greater emphasis on the analysis of purchase documents
- . reliance on protocol analysis
- . time and motion studies
- . input-output analysis
- . reliance on smaller samples

Greater emphasis on the Analysis of Purchase Documents

Examination of organizational buying research, and especially as conducted by industrial firms, suggests a much greater emphasis on secondary data analysis. Of particular interest is the utilization of purchase documents as the data for the analysis of various aspects of organizational buying behavior, such as source loyalty (Wind, 1970). Consumer researchers have relied heavily on panel purchase data (such as MRCA purchase diaries), with little attention, however, to the utilization of household purchase records, e.g., the analysis of bills paid, etc.

Reliance on Protocol Analysis

Protocol analysis—the analysis of transcripts of the verbalized thoughts and actions of a subject when the subject has been instructed to think or solve a problem aloud—has been employed in a number of problem solving situations. The early application of protocol analysis has been in the organizational buying behavior area with work of Clarkson (1963) on trust investment behavior and Wind (1968) on the decision process of purchasing agents. In both cases the protocol approach resulted in considerable insights into the subject's decision process. Consumer researchers should consider the use of this research approach in the study of consumer behavior.

Time and Motion Studies

Time and motion studies have been utilized (although infrequently) to study the work flow of the organizational purchasing function. This approach, whether conducted in the real world (by following consumers around on a shopping trip or during product consumption) or in a laboratory, could be applied to the study of consumer behavior and, in particular, to the study of product (and service) purchase and usage patterns.

Input-Output Analysis

The application of input-output analysis to the organizational buying function offers an interesting way of modeling the purchase function as a provider of inputs for the various users at various stages of the productive process (Watson and Smith, 1966). To date, to the best of this author's knowledge, this approach has not been applied in the study of consumer behavior. It does suggest, however, an approach which could provide intriguing insights into the role of consumer purchases in the maintenance of consumers' activities and standard of life.

Reliance on Smaller Samples

The relatively small number of organizational buyers (which characterize many industries), coupled with the high cost of conducting personal interviews with organizational buyers, have led to frequent reliance on relatively small samples. To achieve effective samples given these constraints, many organizational buying studies rely on careful telephone screening, followed by personal interviews with relatively small samples. Recent difficulties in obtaining consumers' cooperation and the increased cost of consumer interviewing suggest the desirability of following the example of organizational buying behavior studies. Certain advances in segmentation research—componential segmentation—are especially suited for such cases requiring relatively small samples. This approach was developed and first implemented in a number of organizational buying behavior studies (Green, Douglas, and Carmone, 1977). Yet, it is obviously appropriate for consumer segmentation efforts.

Concluding Remarks

The preceding discussion has highlighted some of the potential contributions the organizational buying behavior area can make to the study of consumer behavior. These and other contributions (which time and space constraints prevent me from discussing here) have a number of major implications to the study and teaching of consumer behavior. More specifically, they would suggest that:

1. Consumer researchers could benefit from a careful examination of the organizational buying behavior literature from which they can generate new variables, hypotheses, and even some methodological ideas.
2. The teaching of consumer behavior should not exclude the topic of organizational buying behavior. On the contrary, it seems that the study of organizational buying behavior should be viewed as an integral part of consumer behavior courses.
3. Innovative research on organizational buying behavior is a legitimate area of study for consumer researchers and hence should not be excluded (because of the subject matter) from either the Journal of Consumer Research or the sessions of the Association for Consumer Research.

Accepting my premise and conclusions we are still faced with the question of "how to best proceed toward improving our understanding of which aspects of organizational buying behavior can (and should) be applied to the study of consumer behavior?" Obviously, no prescriptions can, or even should, be given to researchers with respect to style, direction, and approach of their research. Yet, one possible route toward answering this question might be to list the "traditional" differences between organizational and consumer buying behavior and examine the extent to which each of these items can be utilized in the explanation of consumer behavior.

To illustrate this approach, let's consider the dimension of derived demand. The derived demand nature of organizational buying has often been presented as one of the major differences between organizational and consumer buying behavior. Yet, if one accepts the concepts of Becker (1965) and Lancaster (1966) consumer demand for goods can also be seen as a derived demand since goods are in this approach the inputs to the production of "commodities." Recreation, for example, can be considered a commodity which leads to a derived demand for sport equipment and time. Similarly, the demand for health can be viewed as a commodity with a

derived demand for medical care (Grossman, 1972).

These and other approaches should be further explored and experiments designed to test their suggestions concerning the application of organizational buying concepts to the study of consumer behavior. This transfer of organizational buying concepts and methods to the study of consumer behavior is, in my opinion, both feasible and valuable. It offers an exciting new challenge and a legitimate new area of study for consumer researchers.

REFERENCES

- Gary S. Becker, "A Theory of the Allocation to Time," Economic Journal, 75 (September 1965).
- Russell W. Belk, "Situational Variables and Consumer Behavior," Journal of Consumer Research, 2 (December 1975), 157-164.
- Christopher W. Buss, "A Comparison of the Predictive Ability of Group Choice Models," unpublished doctoral dissertation, University of Pennsylvania, preparation.
- Theodore Caplow, Principles of Organization (New York: Harcourt & Brace, 1964), 201-228.
- Jean Marie Choffray and Gary L. Lilien, "An Operational Structure for Assessing Industrial Response to Marketing Strategy," MIT working paper, June 1977.
- G.P.E. Clarkson, "A Model of Trust Investment Behavior," in Cyert and March, The Behavioral Theory of the Firm (Englewood Cliffs, NJ: Prentice-Hall, 1963).
- William M. Evan, "The Organization-Set: Toward a Theory of Interorganizational Relations," in J. Thomason (ed.), Approaches to Organizational Design (Pittsburgh, PA: University of Pittsburgh Press, 1966).
- Amitai Etzioni, Modern Organization (Englewood Cliffs, NJ: Prentice-Hall, 1964).
- Peter Fishburn, The Theory of Social Choice (Princeton: Princeton University Press, 1973).
- Paul E. Green, J. Douglas Carroll, and Frank Carmone, "Design Considerations in Attitude Measurement," in Y. Wind and M. Greenberg (eds.), Moving A Head With Attitude Research (Chicago: AMA, 1977).
- Michael Grossman, The Demand for Health: A Theoretical and Empirical Investigation (New York: NBER Occasional Paper 119, 1972).
- H. Guetzkow and B. Collins, A Social Psychology of Group Processes of Decision Making (New York: Wiley, 1964).
- Jerold Hage and Michael Aiken, Social Change in Complex Organizations (New York: Random House, 1970).
- John K. Hemphill and Charles M. Westie, "The Measurement of Group Dimensions," Journal of Psychology, 29 (1950).
- Terry J. Hillier, "Decision Making in the Corporate Industrial Buying Process," Industrial Marketing Management, 4 (1975), 99-106.
- John A. Howard, Marketing Management: Analysis and Planning (Homewood, IL: R. D. Irwin, 1963).
- _____, Consumer Behavior: Application of Theory (New York: McGraw-Hill, 1977).
- John A. Howard and Jagdish N. Sheth, The Theory of Consumer Behavior (New York: Wiley, 1969).
- Jacob Jacoby, "Consumer Research: Telling It Like It Is," in B. B. Anderson (ed.), Advances in Consumer Research, III, 1976.
- Kelvin J. Lancaster, "A New Approach to Consumer Theory," Journal of Political Economy, 75 (April 1966).
- Sal Levine and Paul White, "Exchange as a Conceptual Framework for the Study of Interorganizational Relationships," Administrative Science Quarterly, 5 (1961).
- George H. Litwin and Robert A. Stringer, Jr., Motivation and Organizational Climate (Cambridge, MA: Harvard Univ. Graduate School of Business Administration, 1968).
- P. K. Patanik, Voting and Collective Choice (Cambridge: Cambridge University Press, 1971).
- James L. Price, Handbook of Organizational Measurement (Lexington, MA: D. C. Heath & Co., 1972).
- P. J. Robinson and C. W. Faris, Industrial Buying and Creative Marketing (Boston: Allyn and Bacon, 1967).
- Jagdish N. Sheth, "Recent Developments in Organizational Buying Behavior," University of Illinois working paper, August 1976.
- _____, "A Model of Industrial Buyer Behavior," Journal of Marketing, 37 (October 1973), 50-56.
- Patricia C. Smith, Lorne M. Kendall, and Charles L. Hulin, The Measurement of Satisfaction in Work and Retirement (Chicago: Rand-McNally, 1969).
- George Strauss, "Tactics of Lateral Relationships: The Purchasing Agent," Administrative Science Quarterly, 4 (September 1962), 161-186.
- C. Watson and W. A. Smith, "Input-Output Analysis and Industrial Purchasing," Journal of Purchasing, 2 (May 1966).
- Frederick E. Webster, Jr., and Yoram Wind, Organizational Buying Behavior (Englewood Cliffs, NJ: Prentice-Hall, 1972).
- Yoram Wind, "The Determinants of Industrial Buyers' Behavior," in Robinson and Faris, op cit., 1967.
- _____, "Applying the Behavioral Theory of the Firm to Industrial Buying Decisions," The Economic and Business Bulletin, 20 (Spring 1968), 22-28.
- _____, "Industrial Source Loyalty," Journal of Marketing Research, 27 (November 1970).
- _____, "A Reward Balance Model of Buying Behavior in Organizations," in G. Fisk (ed.), Essays in Marketing Theory (Boston: Allyn & Bacon, 1971).
- _____, "Preference of Relevant Others and Individual Choice Models," Journal of Consumer Research, 3 (June 1976), 50-57.
- _____, "Organizational Buying Center: A Research Agenda," in G. Zaltman and T. V. Bonoma (eds.), Organizational Buying Behavior (Chicago: AMA, 1977a).
- _____, "Reflections on Creativity and Relevance of Consumer Research," in B. A. Greenberg and D. N. Bellinger (eds.), Contemporary Marketing Thought, proceedings of the 1977b AMA Conference, 55-58.
- _____, "Toward a Change in the Focus of Marketing Analysis: From a Single Brand to an Assortment of Brands," Journal of Marketing (October 1977c).

A COMPARATIVE ANALYSIS
OF DETERMINANT ATTRIBUTES
IN RETAIL STORE SELECTION

Stephen J. Arnold, Queen's University
Sylvia Ma (Student), University of Toronto
Douglas J. Tigert, University of Toronto

Abstract

Results from twelve studies provided the opportunity to identify determinant attributes in retail store selection. Generalizations of the findings appeared possible in light of the variation among the studies in terms of retail environment, market differentiation, and national boundaries. Location and price appeared determinant in food stores selection while value for the money, assortment, and quality were determinant in the selection of fashion clothing stores. The results suggest implications for both super and box store strategies.

Introduction

Review studies such as that of Lindquist (1974-75) attest to the continued search for determinant attributes in the retail store selection process. Is price the major determinant or is it shopping convenience? Or is determinancy dependent upon some combination of other attributes such as assortment, service and quality?

The availability of results from twelve studies provided the opportunity to address these questions. A review of these results is first preceded by a discussion of the concept of determinancy. The results are further placed in context by a brief review of comparable studies. Additional questions are raised and discussed.

Attribute Determinance

The notion of attribute determinance was discussed by Myers and Alpert (1968) and Alpert (1971). In order for an attribute to be determinant in the choice process, or, in other words, predispose consumers to action, the attribute must exhibit two characteristics. First, the attribute must be regarded as important. By important, it is meant that the consumer is "extremely offended" (Myers and Alpert, 1968, citing Foote, 1961) by the attribute's absence. It also appears that "important" is related to the multi-attribute attitude notion (e.g. Fishbein, 1972) that the attribute is "highly evaluated" or that its presence provides the consumers "much satisfaction."

The second characteristic exhibited by a determinant attribute is that alternatives in the choice process are perceived as "differentiated" with respect to the presence of the important attribute. In the multiattribute attitude context, consumers state "different beliefs" or provide "different likelihoods" about the presence or absence of the attribute in each of the alternatives.

According to the preceding, determinant attributes would appear to be identifiable by either of two general approaches. Directly, they can be determined by relating the degree of presence of the attribute to the results of the choice process (Myers and Alpert's, 1968 "covariate analysis"). Less directly, they can be determined by identification of those attributes that consumers say are important, as well as identification of those attributes that consumers say differentiate the alternatives (Myers and Alpert's, 1968 "dual questioning"). Either approach, however, has limitations related to problems of causality and consumer vocalization of "real" motivations and

attitudes (Myers and Alpert, 1968). Following Alpert's (1971) conclusion that regression coefficient determination (covariate analysis) and direct dual questioning are the best methods, it is concluded here that combinations of both general approaches (direct and indirect) are necessary in the search for determinant attributes.

"Important" Attributes in Retail Store Selection

A review by Lindquist (1974-75) of 26 empirical and theoretical studies of retail store selection led him to observe that the following attributes were mentioned in at least 25% of the studies:

<u>Attribute</u>	<u>Scholarly Mentions</u>
Merchandise selection or assortment	42%
Merchandise quality	38
Merchandise pricing	38
Locational convenience	35
Merchandise styling, fashion	27
Service, General	27
Salesclerk service	27

While he provided a caveat, Lindquist suggested that this relative frequency of mention is a "valuable indicator" of the important attributes.

In a study not reviewed by Lindquist, Jolson and Spath (1973) found that price/value relationship, store specialization, quality of merchandise, salesclerk service, and location were the factors considered most important in the selection of eight stores at a local shopping center.

How valid it is to aggregate different retail environments (e.g. food, department, clothing, etc.) in order to identify important attributes is not clear. Presumably, each type of retail store exists to satisfy different consumer needs. Hence, the "evoked set" of attributes from which arise the "determinant" attributes would be different in each situation. Conversely, if all retail environments are characterized by the same evoked set, it would suggest that consumers would be indifferent in their pursuit of need satisfaction as to whether they went to a food, department or clothing store. Casual observation suggests that this indifference is not the case and it is thus concluded that each retail environment must be examined separately.

A separate analysis of department and grocery stores characterized a recent study by Hansen and Deutscher (forthcoming) of important attributes. In this investigation, it was found that the three top-ranked dimensions in grocery store selection were physical facilities (clean, easy to move through, easy to find items, fast checkout), store atmosphere (friendly store personnel), and merchandise (dependable products, high quality products, high value for the money, wide selection, fully stocked, numerous brands, well-known brands, low prices vs. competition, many specially-priced items).

In another recent study of men's clothing stores by James, Durand and Dreves (1976), it was found that price,

assortment and personnel were the attributes deemed most important by respondents.

"Determinant" Attributes in Retail Store Selection

While replication of the preceding studies in similar retail environments might identify consistently-rated "important" attributes, it is not clear that the same subset of attributes would be consistently identified as "determinant." To the extent that markets differ in the extent of differentiation on each of the important attributes, it would be expected that studies carried out in each market would also differ in their reports of determinant attributes. For example, a Toronto, Canada study of retail food stores (Arnold and Tigert, 1973-74) reported that Dominion Store increased its market share from 30% to 45% when it suddenly created a wide gap in the pricing structure by dropping its prices by about 12%. In other words, price appeared determinant among a large group of shoppers.

A year later, there was evidence that the key determinant attribute had changed. With the price war over, "best customer service" was found to be the variable most closely correlated with whether or not the respondent last shopped at Dominion. However, "lower price" still remained most highly correlated with Miracle Food Mart shopping behavior while "easiest to get to" characterized Loblaw's shoppers. That is, even in the same geographic market, determinancy appeared to change over time.

With changes even in the same market, it might be expected that determinant attributes would again be different when national boundaries are crossed and the search for determinant attributes is carried out in different cultures. Thus, Doyle and Fenwick (1974-75) found in a London, U.K. study of food stores that "most individuals preferred to move up the vertical or 'quality' dimension toward Sainsbury's and Marks and Spencer" (as opposed to the "greater variety" or "lower prices" dimensions in that study). Quality is a different determinant attribute than the determinant service, location, and price attributes found in the Toronto study.

It must be further noted that different measurement methodologies characterized the latter two studies. In the Arnold and Tigert (1973-74) study, respondents selected the store best on an attribute. That selection was then correlated against store last shopped. In the Doyle and Fenwick (1974-75) study, however, stores were both ranked by preference and rated on semantic differential scales. These data were then subjected to multi-dimensional scaling. Thus, different methodologies may also account for the different results in these two studies. It is consequently recognized that this factor may tend to confound the search for consistently rated important and determinant attributes.

In view of the preceding considerations, an analysis was made of results obtained in 12 different commercial studies. The usefulness of these studies arose from the fact that although there were differences in methodologies, national boundaries, and retail environments, there were enough replications of each condition to permit tentative conclusions about important and determinant attributes.

Research Methodology

Attribute importance and determinancy in retail store selection were investigated in twelve separate studies characterized by: i) three different countries (U.S., Canada, Netherlands); (ii) three different methodologies (open-ended questions, Likert type statements, forced-choice scales); (iii) three different interviewing techniques (telephone, in-home interview, in-home self-

administered questionnaires); and iv) two different retail environments (supermarkets, fashion). All studies were completed in 1976-77 and all involved random samples of between 1,000 and 3,000 female household heads drawn from Census Metropolitan Trading Areas. All studies involved a measurement of not only attribute importance but also attribute scores of retail outlets, shopping behaviour across outlets, and external validity data on actual prices, market share, performance/productivity, etc. One of the best comparative performance measures across stores and across markets is sales per square foot of gross leasable space (GLA) or sales per square foot of selling area, on either a weekly or annual basis. All ten studies reported here used that measurement, as well as others.

Retail Food

Seven of the twelve studies were surveys of retail food store patronage. In each of these seven studies, the attribute importance question was asked in an identical fashion using the open-ended question, "What is the single most important reason you shop at () for most of your food shopping?" Some 35 specifically recorded verbatim responses were post-coded into eight major components. Six of the seven studies were completed by telephone while the seventh utilized an in-home interview at the same time the telephone interview was being completed in the same market (i.e. the sixth study).

Clothing Fashions

Five separate measures were made of attribute importance in the selection of stores for women's fashions. Two of the studies involved the use of five point importance statements while the remaining three utilized an 11 point forced choice format. Unlike the supermarket studies, the fashion studies involved preselected store attributes that had been determined on the basis of a large number of focussed group interviews probing on the importance of various store characteristics. The derived list of 11 attributes included a number of economy/utility dimensions (price, value, location) as well as a number of fashion specific store characteristics. Three of these dimensions relate to how a store is positioned on the "fashion spectrum" from low (everyday, conservative wear) to high (latest, most fashionable women's wear). Respondents completed the questionnaires in their homes on a self-administered basis.

Results

Retail Food

Table 1 reports on the seven separate studies of supermarket shopping behavior. While the commercial nature of the data dictates disguise of each city's identity, it is nonetheless clear that location and price ranked either first or second in importance among the eight major attribute groupings.

With regards to determinancy, it was observed that in all seven studies, the chain with the highest market performance, as measured by sales per square foot of GLA, was also the market price leader, both in terms of consumer perceptions and in terms of an actual market basket measurement, or was within two percentage points of the price leader in the shopping basket measure. Price was thus concluded to be determinant.

There are several differences in the columns of Table 1 that relate primarily to differences in competitive activity and hence market differentiation but which reinforce the above conclusions about importance and determinancy.

TABLE 1

COMPARATIVE ATTRIBUTE IMPORTANCE DATA FOR SUPERMARKETS: SHARE OF MENTIONS GOING TO EACH RESPONSE TO THE QUESTION..."ALL THINGS CONSIDERED, WHAT IS THE SINGLE MOST IMPORTANT REASON YOU SHOP AT (_____) FOR MOST OF YOUR FOOD SHOPPING"

REASONS GIVEN*	CITIES						
	NORTH AMERICA				NETHERLANDS		
	A	B	C	D	E	TEL	IN-HOME
1. Location/convenience/easy to get to.....	34%**	35%	36%	34%	27%	23%	22%
2. Prices/low price.....	14	18	16	13	40	33	37
3. Large variety/well stocked/assortment/everything you need.....	9	11	9	11	7	13	12
4. Meat quality/variety.....	11	6	8	7	5	0	0
5. Overall quality.....	9	3	5	5	1	6	5
6. Service/friendly/courteous/fast checkout.....	4	6	7	6	5	7	5
7. Specials/sales, coupons/stamps/weekly specials.....	7	5	6	7	3	2	1
8. Pleasant shopping environment.....	0	0	1	2	0	8	7
9. All others.....	<u>12</u>	<u>16</u>	<u>12</u>	<u>15</u>	<u>12</u>	<u>8</u>	<u>11</u>
TOTALS.....	100%	100%	100%	100%	100%	100%	100%
Sample Size (N).....	(1000)	(1000)	(1000)	(1000)	(1000)	(1000)	(1000)

* Question was asked in open-end form. Responses were recorded verbatim and later post-coded into the above categories.

** In U.S. city "A", 34 percent of the total sample said location/convenience was the most important reason why they chose the store where they do most of their food spending.

The first four cities (A, B, C, D) all show location/convenience to rank first by a wide margin with low price in second position but only slightly ahead of attributes like assortment, meat, quality and service. In these four markets, the price differential across the four leading chains in market share is less than four percentage points in the price basket. Thus, while there is limited but noticeable price differentiation, price clearly plays a lesser role as a determinant of patronage and the non-price variables play a more prominent role, particularly location.

In North American city "E" and in the two Netherlands studies, the low price attribute jumps into first place and by a wide margin over location. In these markets, the price leader is 12 to 15 percent below the next nearest competitor. Furthermore, the price leaders are also the market share leaders in these markets and have by far the highest performance records in sales per square foot. Clearly, the competitive structure of the market has forced price into the leading role as a determinant store attribute. In addition, but not shown in this report, low price is the single most important determinant of patronage for the market leaders in these markets. Finally, on store attributes like cleanliness, friendly personnel, etc., that received high ratings in the Hansen and Deutscher study, the leading chains in city "E" and the Netherlands cities have the lowest ratings in the market. They have traded away those

dimensions, as well as quality on such items as meat and produce, in favour of total market dominance on price and assortment. Their bottom line profitability indicates their strategy is working well.

In the two Netherlands cities, meat quality/variety received no mentions on the attribute importance question. That result should hardly be surprising given that upwards of 80 percent of all Dutch consumers buy their meat at a specialty butcher shop. Pleasant shopping environment, however, ranks fourth in the Netherlands while receiving almost no mentions in the North American studies.

Clothing Fashions

The results of the comparative investigation of important attributes in the selection of women's fashion clothing stores are found in Table 2. Value for the money, assortment/selection, and quality dominate the first three rank positions. Rank correlations above 0.99 for all 10 pairs of columns indicate that these results are highly consistent across the five studies.

Across national boundaries and assuming equivalent translation of the questions, it was found that there were significant differences in the mean score levels. For example, in the two forced choice scales in Canada in 1976 and 1977, "best value for the money" totally

TABLE 2

COMPARATIVE ATTRIBUTE IMPORTANCE DATA FOR WOMEN'S CLOTHING FASHIONS: MEAN AND RANK SCORES ON STORE CHARACTERISTICS FOR LIKERT SCALES AND FORCED CHOICE SCALES..."HOW IMPORTANT IS EACH SPECIFIC STORE CHARACTERISTIC TO YOU IN CHOOSING A STORE TO SHOP FOR WOMEN'S FASHIONS?"

STORE CHARACTERISTIC	5 POINT LIKERT SCALE				11 POINT FORCED CHOICE SCALE					
	NETHERLANDS (77)		CANADA (76)		NETHERLANDS (77)		CANADA			
	Mean	Rank	Mean	Rank	Mean	Rank	1976	1977	1976	1977
1. Gives best value for the money.....	4.5	1*	4.6**	1	3.9	2	2.5***	2.6	1	1
2. Has largest overall assortment/ selection.....	4.4	3	3.9	2	4.4	4	4.5	4.1	2	2
3. Has the highest quality women's fashions.....	4.5	2	3.8	3	3.2	1	4.9	5.5	3	4
4. Has the most knowledgeable, helpful salesclerks.....	4.2	4	3.6	4	4.1	3	5.3	6.1	4	6
5. Best for current, up-to-date women's fashions.....	3.9	5	3.6	5	6.0	7	5.4	5.2	5	3
6. Easiest to get to from home.....	3.6	6	3.1	7	4.9	5	5.6	5.8	6	5
7. Best for conservative, everyday wear.....	3.5	8	3.5	6	6.4	8	6.0	6.5	7	7
8. Best for latest, most fashionable women's wear.....	3.5	7	3.0	8	6.9	9	7.2	6.9	9	8
9. Has the lowest prices.....	3.2	9	3.0	9	5.8	6	6.3	7.0	8	9
10. Has the best fashion advertising..	2.6	10	2.4	11	9.2	11	8.8	8.8	11	11
11. Has the most exciting merchandise display.....	2.6	11	2.9	10	9.0	10	7.8	7.4	10	10

* Rank scores involve examination of mean scores taken to third decimal place. For ease in reading, only first decimal is shown in the above table.

** Five point Likert scale ranges from "very important" (score of 5) to "not important at all" (score 1) with highest mean score ranking first.

*** Eleven point forced-choice scale ranges from "most important" (score 1) to "least important" (score 11) with the lowest mean score ranking first.

dominated the data with mean scores of 2.5 and 2.6 in the two years. The gap between this dimension in first place and the second ranked dimension, largest overall assortment, is enormous, about 20 standard errors. In the Netherlands, the top three store dimensions received almost identical mean scores on the importance scale and "highest quality" was in first place on the forced choice. In fact, highest quality ranked first or second in both Dutch studies.

An interesting trend was also observed in the two year period separating the two Canadian studies. "Best for current, up-to-date women's fashions" moved up from fifth to third position while "knowledgeable, helpful salesclerks" dropped from fourth to sixth position. While such a change could be attributed to sampling error, focussed group research suggested that salesclerk service was becoming universally poor across major fashion outlets with the result that the consumers perceived they would have to service themselves in the stores. Thus, importance of good in-store service was consequently declining as an important store attribute.

Discussion and Implications

What conclusions can be reached on the basis of these results and in the context of other studies devoted to the identification of important and determinant attributes? First, price/value for the money is an important attribute and will be determinant, it appears, with even small amounts of differentiation across competitors.

This finding, characterizing both the retail food and fashion clothing environments, clearly matches the results of Jolson and Spath (1973) and James *et al* (1976) and closely parallels those of Lindquist (1974-75) and Arnold and Tigert (1973-74). Only in the results of Hansen and Deutscher (forthcoming) does the effect of price appear absent. It is hypothesized that the latter result was obtained because there was little pricing differentiation in the Columbus, Ohio market at the time of their study and that respondents to the questionnaire were, in fact, identifying the determinant as opposed to the important attributes as suggested. Thus, overall, the economic model of the consumer is suggested as having strong validity in the retail store selection process.

Attributes either equally important or secondary to price appear to differ depending upon the environment. Among retail food stores, according to these studies, it is locational convenience. Among fashion clothing stores, however, it is assortment and quality.

In other words, there is considerable evidence that whenever there is noticeable dispersion across supermarket chains on the economy/utility dimensions (price, location), those store characteristics dominate the store choice process. Fashion shoppers, on the other hand are not as concerned about physical convenience. They search out value, quality, wide assortments, and up-to-date fashions. They appear to be willing to travel to whatever location is required to find what they want.

This second conclusion supports Arnold and Tigert (1973-74) in the retail food environment and James et al (1976), Lindquist (1974-75), and Jolson and Spath (1973) to the extent that their findings apply to fashion clothing stores. These two conclusions appear to hold implications for both retail food superstores and box stores. The supermarket chains that are now moving into the "superstore" game with 35-100,000 square foot outlets are implicitly trading on the assortment attribute which was not concluded to be as important as price and location. Thus, in order to be successful, superstores will at least have to match the competition on price. But the very nature of the large capital investment in physical plant, however, implies fewer stores and therefore less convenience. Their only apparent option in order to satisfy the important convenience attribute would be, as some have done, to add specialty bakeries, delicatessens, floral shop and general merchandise departments, which would then provide one-stop shopping for the consumer and thereby justify the greater distance they would have to travel on average. Superstores would then be perceived as satisfying the convenience criterion.

At the other end of the spectrum are the box stores such as Aldi in Iowa which are attempting to take a share of the market on the basis of only limited assortments of basic groceries (400 stock keeping units). Furthermore, prices are at least ten percent below the next nearest competitor.

While the box store strategy has worked well in Germany, the evidence is not yet available for North America. However, the results of this study appear to support the strategy. Pricing well below the competition automatically makes the important price attribute also determinant. In addition, low capital investment per outlet should also make them more amenable to greater numbers of units and hence more likely to satisfy the convenience attribute.

If the future brings a showdown between the superstore and the box store, the superstore may win. It is conceivable that over time, the cost advantages that accrue to the box store will be wiped out due to advances in technology. For example, no price stamping of merchandise in the box store could be countered by optical scanning of label codes. Thus, if price differentiation is erased and consumers perceive the fewer superstores just as convenient, the superstores will then completely dominate on assortment, the other important and consequently determinant attribute in retail food store selection.

References

- Mark I. Alpert, "Identification of Determinant Attributes: A Comparison of Methods," Journal of Marketing Research, 8(May, 1971), 184-91.
- Stephen J. Arnold and Douglas J. Tigert, "Market Monitoring Through Attitude Research," Journal of Retailing, 49(Winter, 1973-74), 3-22.
- Peter Doyle and Ian Fenwick, "How Store Image Affects Shopping Habits in Grocery Stores," Journal of Retailing, 50(Winter, 1974-75), 39-52.
- Martin Fishbein, "The Search for Attitudinal-Behavioral Consistency," in Joel B. Cohen (ed.), Behavioral Science Foundations of Consumer Behavior (New York: Free Press, 1972).
- Nelson N. Foote, Consumer Behavior: Household Decision-Making, Vol. 4 (New York: New York University Press, 1961).

Robert A. Hansen and Terry Deutscher, "An Empirical Investigation of Attribute Importance in Retail Store Selection," Journal of Retailing, forthcoming.

Don L. James, Richard M. Durand and Robert A. Dreves, "The Use of Multi-Attribute Attitude Models in a Store Image Study," Journal of Retailing, 52(Summer, 1976), 23-32.

M.A. Jolson and Walter F. Spath, "Understanding and Fulfilling Shopper's Requirement: An Anomaly in Retailing?" Journal of Retailing, 49(Summer, 1973), 39-49.

Jay D. Lindquist, "Meaning of Image: A Survey of Empirical Hypothetical Evidence," Journal of Retailing, 50(Winter, 1974-75), 29-38.

James H. Myers and Mark I. Alpert, "Determinant Buying Attitudes: Meaning and Measurement," Journal of Marketing, 32(October, 1968), 13-20.

A PRAGMATIC APPROACH FOR RETAIL FASHION MONITORING

Lawrence J. Ring, University of Virginia

Abstract

The need to monitor market changes, competitive activities, and individual company performance is a challenge to all retailers. And, in the fashion retailing business, it is especially important because of frequently changing consumer tastes and interests and the attendant shifts in merchandise styles and merchandising methods. This paper outlines a method for fashion market monitoring which tracks both consumer interests and store patronage and can be used to better understand store performance.

Introduction

Within the past decade, the fashion arena has become a popular topic of contemporary consumer research. Over the same period, the proliferation of new styles and fashions has pushed the industry to new sales heights here and abroad. Recent figures indicate that U.S. and Canadian consumers purchase well over 25 billion dollars worth of clothing and accessories annually from a diverse array of retail outlets.

In this mass market, recent studies (King, Tigert, and Ring, 1975, and Sproles and King, 1973) reveal that a large proportion of the population is interested in fashion participation and continuously monitors current fashion trends. In addition, mass fashions are simultaneously available to the fashion consumer in nearly all price ranges.

Mass communications and the rapid production of fashions have increased the rate of fashion diffusion throughout the population. We are now witnessing the accumulation of sizable and frequently updated wardrobes by almost all segments of the population.

Along with this development of mass manufacturing strategies, the proliferation of mass media information and influence, and the increased levels of disposable income has come a general explosion of new retailing and merchandising concepts which have strong implications for the individual firm's merchandising strategy and attendant "retail image."

In any particular fashion retail marketplace, the consumer is faced with a choice among a varied group of fashion outlets each developing and maintaining its own "fashion" profile. The ability of each of those competing retailers, and particularly among those large chains competing for the mass market, to meet consumer needs and at the same time monitor competitive changes in the marketplace is dependent upon the individual firm's ability to efficiently gather and synthesize information about the changing marketplace.

This paper describes a methodology for providing a complete diagnosis of the structure of an individual fashion market including patterns of buying behavior and determinants of patronage for major competing chains.

Fashion Involvement

Historically contemporary fashion research has focused on the fashion innovator and the fashion opinion leader

as prime sales targets and key links to the volume fashion market. A variety¹ of methodological studies designed to define, isolate, and profile these segments have dominated fashion research since the basic concept was first introduced by King (1963).

In general, fashion segmentation has been based on the division of consumers into relatively homogeneous segments based on demographic, psychological, and sociological measures, media exposure characteristics, and specific measures of fashion involvement and consciousness.

Contemporary fashion theory research has tended to focus on the fashion process and its interacting behavioral system. Researchers have been increasingly interested in the dynamics of fashion adoption and diffusion. The specific situations of fashion innovativeness, fashion opinion leadership, and the broad concept of the fashion change agent provide an expanding base of literature in the fashion arena.

In fashion apparel thought, the fashion change agent has been seen as more than simply an innovator or opinion leader as traditionally defined. Sproles and King (1973) have argued that while innovativeness and opinion leadership in fashion are theoretically distinct concepts, that the differentiation may not be either pragmatically necessary nor even a requirement for segmentation purposes. The time compression of contemporary fashion diffusion confounds the strict identification of innovators and opinion leaders and major or mass market adoption is accelerated instead by a broadly identified change agent segment of the population (Sproles and King, 1973).

From the retailer's point of view in the mass market, the problem is to determine the levels of fashion involvement of the population and ascertain the size and needs of each "level of fashion involvement" segment.

Fashion Involvement Measures

Over the last fifteen years, the author and others have conducted a number of research investigations of fashion involvement in the United States and Canada. The accumulated data base now contains measurements of fashion interest at different points in time in different geographic regions and among both male and female populations. As a result, a comparative data base spanning fifteen years has now been established.

In the case of one measure, data has been collected from over twenty independent populations using essentially identical methods of sample selection and exactly identical and, therefore, comparable measurements of fashion interest. That measure is general fashion interest (also sometimes referred to as fashion awareness). In

¹A representative list of fashion segmentation studies would include King (1963), Grindereing (1967), Pasnak and Ayers (1969), King and Summers (1970), Reynolds and Darden (1973), Sproles and King (1973), Baumgarten (1975), Tigert, Ring, and King (1975), and Painter and Grianzin (1976).

each of the investigations mentioned above, the following question was asked of all respondents:

"Which ONE of these statements best describes your reactions to changing fashions in men's (women's) clothes? Even though there may be no statement listed which exactly describes how you feel, make the best choice you can from the answers listed."

I read the fashion news regularly and try to keep my wardrobe up to date with the fashion trends . . . 1

I keep up to date on all the fashion changes, although I don't always attempt to dress according to these changes 2

I check to see what is currently fashionable only when I need to buy some new clothes 3

I don't pay much attention to fashion trends unless a major change takes place 4

I am not at all interested in fashion trends. . . 5

Comparative results using this measure have been reported elsewhere and provide a baseline for future comparison and tracking of fashion interest (see King, Tigert, and Ring, 1975, and Ring, 1977). Therefore, this particular question provides an excellent measure of fashion interest across any particular population.

Attitudes Toward Stores/Determinants of Patronage

The relationship between consumer attitudes and retail sales has long been a popular topic for examination by retail researchers -- and particularly with regard to department stores (Berry, 1969, Lazer and Wyckham, 1969, and Weale, 1971, for example.)

A number of techniques have been utilized by marketers in researching retail image. The semantic differential (Arons, 1961, Anderson and Scott, 1970, and Kelly and Stephenson, 1967) and the open-ended questionnaire (Berry, 1969), have been two of the more popular techniques employed.

As discussed by Arnold and Tigert (1973), however, each of these techniques has significant problems. The major issue is their characteristic length and complexity. The semantic differential, for example, requires the respondent to rate each retailer on each attribute under study. The open-ended questionnaire, on the other hand, generally requires personal interviewing and content analysis of the responses.

Tigert (1976) has demonstrated that the measurement of store characteristics by the use of the "associative technique" is both reliable and valid, as well as discriminating. He contends, in fact, that the associative technique measures the world in a way which reflects the manner in which consumers go through the choice process.

Operationally, the associative technique involves asking each respondent which store (among a set of stores representative of a particular market) best answers each of a set of individual questions related to the stores. This technique has been employed in the analysis which follows.

Patronage Measures

Retail image researchers have, over the years, identified a large number of dimensions upon which consumers judge a store's performance and which therefore may conceivably contribute to the constitution of a

store's personality or image. For example, Lazer and Wyckham (1969), listed four major image dimensions, while Martineau (1958) suggested six traits, Stephenson (1969) described eight dimensions, Wyckham (1967) 30 dimensions, May (1971) 42 dimensions, Berry (1969) 44 dimensions and Shapiro (1970) 94 dimensions in a study of retail products among Boston consumers.

While these studies have focused on a variety of store types and product classes, many have dealt in some way with fashions. Based on these results, and on other fashion research over a period of several years, several "attitudes" or "image dimensions" have been found to consistently characterize retail fashion chains and as well to be useful for management decision-making. These dimensions in the question format of the "associative technique" and in the case of men's fashions are:

1. Which store is the easiest one to get to from your home?
2. Which store has the lowest prices?
3. Which store has the highest prices?
4. Which store has the most knowledgeable/helpful salesclerks?
5. Which store has the highest quality men's fashions?
6. Which store has the lowest quality men's fashions?
7. Which store gives you the best value for the money in men's fashions?
8. Which store gives you the worst value for the money in men's fashions?
9. Which store has the largest overall merchandise assortment/selection in men's fashions?
10. Which store is best for conservative, everyday men's wear?
11. Which store is best for current, up-to-date men's wear?
12. Which store is the best for the very latest, most fashionable men's wear?
13. At which store do you find the most exciting merchandise display of men's fashions?
14. Which store has the best fashion advertising?

In addition, two measures of share of shoppers should be employed in this approach, again using the "associative technique." They are:

1. Which store was visited the last time an article of men's wear was purchased?
2. Which store is shopped most often for men's wear?

Toronto: An Example

The results which are reported later in this paper represent an example of the implementation of the approach described here over several years in Toronto, Ontario, Canada for the expressed purpose of monitoring the retail fashion marketplace there. This research was conceptualized as a joint academic/industry project and has been operated and supported on that basis now for four consecutive years beginning in 1974. The example which is discussed below is based on data collected about the men's fashion market in Toronto and is repre-

sentative of other similar efforts elsewhere and in the women's fashion market as well.

Data Collection

The data for the Toronto Project has been collected each year (1974-1977) by mailed self-administered questionnaire from the Toronto Census Metropolitan Area (CMA). The samples have been obtained by telephone placement of questionnaires to adult male household heads. The samples were recruited to be in balance with city demographic profiles. Questionnaire administration has taken place in the spring of each year. A total of 615 usable questionnaires were mailed to each respondent complete with standard self-administration instructions. Each participant received a small gift for completing the questionnaire.

Store Selection

In attempting to understand the structure of the Toronto men's wear retail mass marketplace, it was first necessary to define a relevant set of retailers to present to respondents for their judgments. Since the analysis was to be concerned with the mass market, and was to be based on consumer judgment, it was important to include those chains which had sufficient market coverage and general recognition for respondents to judge them.

Ten stores and an "all other" category were listed as response alternatives in the survey, including the city's three major department stores. Each of the chains included in the analysis was operating at least six retail outlets in Toronto in 1975. In addition, it was estimated that each of the chains was obtaining gross men's wear sales in Toronto of at least \$5 million per year.

Fashion Interest Results

Table 1 presents the 1974-1977 fashion interest results for the ongoing Toronto project. These results can be interpreted in several ways. For example, those respondents

who perceived themselves to be in the first category (see question series), might be identified as innovators since they indicate that they keep their wardrobes up to date with fashion trends. However, perhaps a more important concern than the first category alone is the aggregated total of the first two categories. Pragmatically, those respondents who have responded in either the first or second level fashion interest categories might be classified as active in fashion interest, at least in terms of consistently monitoring, if not purchasing, new fashions.

By combining the first and second levels of fashion interest as indicated above, it is clear that approximately one quarter of the respondents over the four years can be classified as active in fashion interest and capable of influencing the mass population at least verbally, if not visually, with regard to changing fashion trends. This is true because those men responding in both the top two fashion interest levels keep informed of fashion changes and thereby could potentially be communicators of information. And, those individuals in fashion interest level one are also potentially visual influentials by virtue of the fact that they keep their wardrobes up to date and are therefore wearing the latest fashions for all to see.

Looking at the data another way, it appears that the majority of adult male Toronto consumers are not particularly interested in new men's fashions. And, in comparison to female and college student fashion interest statistics and even American adult male statistics (see Ring, 1977), Toronto has fewer change agents (categories one and two).

More importantly, from a market monitoring perspective, it appears some significant shifts have occurred in the level of fashion interest over the four year data base. For example, the drop in percentage reporting as fashion change agents from 1974 to 1975 represented a significant change in fashion interest level, as did the five percent increase from 1976 to 1977.² This kind of information might be particularly important to the firm which is mer-

TABLE 1
TORONTO MALE FASHION INTEREST

Statements Describing Degree of Fashion Interest ^a	1974 Toronto	1975 Toronto	1976 Toronto	1977 Toronto	Average
Read fashion news regularly and keep wardrobe up-to-date with fashion trends	4%	2%	1%	2%	2.25%
Keep informed of fashion changes but do not always follow	21	16	15	19	17.75
Check what is fashionable only if buying new clothes	24	28	28	30	27.5
Only pay attention to major fashion changes	26	30	36	33	31.25
Not at all interested in fashion	25	24	20	16	21.25
TOTAL NUMBER RESPONDING	100% (615)	100% (1,017)	100% (865)	100% (983)	100%

^aIdentical measures of interest used in all surveys.

*READ: In 1977, 2% of the sample said they read the fashion news regularly, and kept their wardrobes up-to-date with the fashion trends.

²The observed difference of 7 percent is greater than 2.4 standard errors from zero, while the 5 percent difference from 1976 to 1977 is 1.96 standard errors from zero. Both significant at the .95 level.

chandising primarily toward the top end of the fashion market as is demonstrated in the next section of this paper.

Fashion Interest by Store

The basic thesis in this analysis is that consumers differ in the fashion shopping behavior based on their self perceived fashion interest or involvement. To explore this question, the fashion interest measure need only be cross-classified with a measure of share of shoppers such as "store shopped most often."

The following example illustrates some rather dramatic differences between the self-perceived fashion interest level of the shoppers of a leading Toronto department store and a leading fashion specialty chain. In Table 2, it is clear that the fashion specialty chain shoppers perceive themselves as much more interested in men's fashions than do the department store shoppers. Taking 1975 as an example, fully 40 percent (levels 1 and 2) of the fashion specialty chain shoppers were regularly monitoring (if not purchasing) new men's fashions. In contrast, only 17 percent of the department store shoppers made a similar response.

These differences are important at any given point in time and also as they may change over time to the management of each chain. Satisfying the needs of the highly fashion interested clientele of the fashion specialty chain surely points to a different merchandising strategy than does satisfying the needs of the department store shopper. These results lead to an analysis of store patronage dimensions.

TABLE 2
TORONTO 1975
CROSS CLASSIFICATION ANALYSIS:¹
"CHAIN SHOPPED MOST OFTEN" vs. "FASHION INTEREST"

Fashion Interest: Reaction to Changing Fashion	Chain Shopped Most Often	
	Department Store	Fashion Specialty Chain
1. Read fashion news regularly and keep wardrobe up-to-date with fashion trends.	0.8%*	4.4%
2. Keep informed of fashion changes but do not always follow.	16.9	35.6
3. Check what is fashionable only if buying new clothes.	28.5	35.6
4. Only pay attention to major fashion changes.	34.1	22.2
5. Not at all interested in fashion.	19.7 100.0%	2.2 100.0%

¹The difference between department store consumer fashion interest and fashion specialty chain fashion interest were significant at the .001 level based on X² analysis.

*READ: Among those who shop most often at the department store, 0.8 percent said that they regularly read the fashion news and tried to keep their wardrobes up-to-date with the fashion trends.

Determinants of Patronage

The analysis of store patronage can proceed at two levels. Initially, the concern is with the share of shoppers mentioning each chain on each of the patronage dimensions. Secondly, however, it is important analyze the attitudes of each chain's shoppers on a chain-by-chain basis.

Table 3 continues the example of the department store and the fashion specialty chain for the year 1975. In this table the share of shoppers (of the overall sample) which mentioned each of the two stores as being best (or worst, highest, lowest, etc.) on each of the patronage dimensions have been presented. These results document the "image" of each chain as seen by the overall sample at a particular point in time (spring, 1975.)

TABLE 3
SHARE OF MENTIONS FOR TWO CHAINS
ON THE DETERMINANTS OF PATRONAGE QUESTIONS

Determinant of Patronage	Chain Mentioned	
	Department Store	Specialty Chain
1. Easiest to get to from home	18%*	2%
2. Store Last Shopped	29	5
3. Store Shopped Most Often	35	6
4. Lowest Prices	7	0
5. Highest Prices	3	48
6. Highest Quality	9	41
7. Lowest Quality	1	0
8. Best Value for the Money	28	4
9. Worst Value for the Money	3	7
10. Most Knowledgeable, Helpful Salesclerks	19	20
11. Largest Overall Assortment/Selection	30	6
12. Best Advertising	30	9
13. Most Exciting Display	14	22
14. Best for Conservative Everyday Men's Wear	31	5
15. Best for Current Up-to-Date Men's Wear	12	23
16. Best for Very Latest Most Fashionable Men's Wear	7	29

*READ: Eighteen percent of the total 1975 sample said that the department store was the easiest to get to from home, while only 2 percent said that the fashion specialty chain was the easiest to get to from home; 29 percent of the sample said that they last shopped at the department store, while 5 percent said that they last shopped at the specialty chain.

Market coverage in Table 3 has been measured by asking respondents which store is the "easiest to get to from their home?" Based on the percentage of respondents who identify a particular store as the easiest on to get to from home, a "trading area" can be defined for each store (chain). The discussion will refer to the trading area with the term, "market coverage." For example, the Department Store had a 18 percent market coverage because 18 percent of the respondents said that the Department Store was the easiest to get to from home. The Specialty Chain had only a 2 percent market coverage.

A more important concern than market coverage alone is the relationship between market coverage and share of shoppers (as measured by the "last shopped" and "shop most often" questions.) Twenty-nine percent of the sample last shopped at the department store compared to 5 percent which last shopped at the fashion specialty chain. Similar figures are obtained on the shop most often measure, indicating that a much larger percentage of the market shops at the department store than at the specialty chain.

However, the ratios between "market coverage" and the two share measures provide a slightly different perspective. The ratio analysis eliminates the differential market coverage across chains and expresses each chain's customer set as a ratio of coverage.

		Ratios	
		<u>Department Store</u>	<u>Specialty Chain</u>
1.	<u>Last Shopped</u> Market Coverage	1.6/1	2.5/1
2.	<u>Shop Most Often</u> Market Coverage	1.9/1	3.0/1
3.	<u>Lowest Prices</u> Highest Prices	2.3/1	0.0/1
4.	<u>Highest Quality</u> Lowest Quality	9.0/1	41.0/0
5.	<u>Best Value</u> Worst Value	9.3/1	0.6/1

Both of the chains in the above example have better than "one to one" ratios of the two share measures to market coverage. Interpretatively, these findings indicate that these two chains were performing better than could be expected based on their market coverage alone. Chains with ratios of less than the "one to one" would be drawing fewer shoppers than the percent who said the chains were closest to home--meaning some shoppers were driving past those chains to get to other, less accessible stores for some reason.

The ratio analysis can also be performed on the price, quality and value for the money measures since two data points were collected for each of those dimensions. A better than "one to one" ratio on each of these dimensions would indicate a positive price, quality, and value image from the overall sample. In this example, the department store is positive on all three dimensions, while the fashion specialty chain is positive on quality and "negative (or less than one to one) on price and value.

Looking at some of the other dimensions, the department store appears to have a larger percentage than the specialty chain in many, but not all, of the cases. This may, in part, be attributable to the fact that many more respondents shop at the department store, and all other things being equal probably tends to rate the store

shopped most often highest or best on each of the important dimensions. Consequently, to some extent the share of mentions on any given question often tends to mirror the share of shoppers for a given chain. Thus, if the results are examined only in terms of absolute percentages, a store with a large share of shoppers might be thought to be very strong on a particular dimension, while a store with a small share might be considered to be weak. This mistake can be avoided if the raw percentages are examined in comparison to the overall share of shoppers for the chain. For example, in terms of absolute percentages, both of the example chains appear to be equally strong on the best salesclerk question (question 10.) The department store obtains 19 percent of the mentions and the specialty chain obtains 20 percent. However, if these figures are compared to overall share of shoppers, it is clear that the department store is not performing particularly well on this dimension, while the specialty chain is extremely strong here. Only 6 percent of the sample shopped at the specialty chain but three times that many said the chain had the most knowledgeable, helpful salesclerks. In contrast, 35 percent of the sample shopped at the department store, but only 19 percent said that store had the most knowledgeable, helpful salesclerks.

Another method of understanding the strengths and weaknesses of each chain is to analyze how the shoppers of each individual chain rate that chain. This can be easily accomplished through the use of cross-classification analysis.

Cross Classification

Table 4 contrasts the shoppers of the two chains across 11 determinants of patronage. This analysis is based on the cross-classification of shoppers "most often shopping" each chain with each of the retail patronage dimensions. Given the fashion interest findings discussed earlier, the fashion specialty chain should be expected to be perceived as more fashion-oriented than the department store. In this example, that is the case.

Clearly, the shoppers of the department store have a different "image" of their "most often shopped" chain than do the fashion specialty chain shoppers of their "most often shopped" chain.

Although more sophisticated analysis techniques can be utilized (for example, see Ring, 1976), simply assigning the attributes to whichever chain obtained the higher percentage yields interesting results. What emerges is a statement of the strengths and weaknesses of the two chains as perceived by their own shoppers.

For example, consumers would appear to be shopping at the department store because of its good location, lower prices and higher value for the money, better merchandise assortment/selection, better (more) advertising, and a stronger orientation to conservative, everyday men's wear. The fashion specialty chain shoppers, on the other hand, are paying for higher quality, more salesclerk "stroking", an exciting merchandise display, and a current-to-very latest, most fashionable apparel mix.

These findings are consistent with the consumer self-perceptions of their own fashion interest as discussed previously. The department store was viewed as low on the fashion spectrum and its shoppers perceived themselves as only marginally fashion involved. The fashion specialty chain shoppers perceived themselves as much more highly interested in fashion and their chain as highly fashion oriented in terms of product offering.

TABLE 4
TORONTO 1975
CROSS CLASSIFICATION ANALYSIS:
SUMMARY OF DIAGONALS FOR DETERMINANTS OF PATRONAGE

Determinants of Patronage	Chain Shopped Most Often	
	Department Store	Specialty Chain
1. Easiest to Get to from Home	34%*	24%
2. Lowest Prices	17	0
3. Highest Quality	20	91
4. Best Value for the Money	64	40
5. Most Knowledgeable, Helpful Salesclerks	44	85
6. Largest Overall Assortment/Selection	58	27
7. Most Exciting Merchandise Display	31	86
8. Best Advertising	51	33
9. Best for Conservative, Everyday Men's Wear	61	23
10. Best For Current, Up-to-Date Men's Wear	27	75
11. Best for Very Latest, Most Fashionable Men's Wear	15	77

*READ: Among those who shopped most often at the department store, 34 percent said that the department store was easiest to get to from home, 17 percent said the department store had the lowest prices, and 20 percent said the department store had the highest quality, etc.

Monitoring Over Time

Thus far the discussion of store patronage and the fashion profiles of each store's shoppers has been discussed only as it related to one point in time, the spring of 1975. While this analysis in and of itself is obviously useful, it becomes even more relevant when it is done on a regular basis over a long period of time. The time series analysis permits management to observe its effectiveness at attempting to penetrate new market segments or to change its image from year to year. For example, one major Toronto department store had watched its market share and its image on many important dimensions such as display, assortment, and merchandise fashionability erode as competitors expanded their operations into new stores with more fashionable merchandise while the department store "let itself go" in anticipation of the eventual opening of a major downtown showcase flagship store and shopping mall. The data in Table 5 reflect this store's performance on selected dimensions over the four year period during which the Toronto project has operated.

Between the collection of data in 1976 and in 1977, this department store opened its new flagship store downtown. The upward shifts on the selected dimensions included in the table reflect the impact of the new store on the marketplace, and indicate the chain's image has improved over the year on each of these measures.

TABLE 5
SELECTED DETERMINANTS OF PATRONAGE FOR A
TORONTO DEPARTMENT STORE, 1974-1977

Determinant of Patronage	1974	1975	1976	1977
	1. Largest Assortment	34%*	29%	24%
2. Most Exciting Display	21	15	12	23
3. Best for Current Men's Wear	14	9	8	13
4. Best for Very Latest Men's Wear	13	7	6	12

*READ: In 1974, 34 percent of the respondents said this department store had the largest assortment, while only 29 percent made the same statement in 1975, 24 percent in 1976, and then back up to 28 percent in 1977.

Similar analyses have been completed for each chain on each of the patronage dimensions and on the fashion interest question for each of the four years. In addition, a number of other relevant dimensions have been measured as well, included demographics, fashion AIO measures, price and volume pointing measures, additional fashion involvement questions, media and information sources measures, and fashion item (suit, sports jacket, etc.) measures. In total, these results provide the basis for a complete diagnosis of the Toronto fashion market including patterns of buying behavior and determinants of patronage for major competing chains.

Summary

Earlier studies have documented major differences among consumers in terms of their fashion involvement and fashion consciousness. Other research has uncovered major differences in consumer perceptions of competing retailers and in the profiles of the customers of competing retailers

In the competitive fashion retailing milieu, these differences are of major decision-making importance. The established retailer inevitably has a store image which has attracted his existing clientele. That image may have been planned as a strategic maneuver, or it may have been generated fortuitously. Therefore, the retailer, in selecting his product line offerings, as well as in designing his entire fashion retail presentation, can buy product offerings which coincide with his present strategy and image, or select offerings or make image decisions purposively designed to attract other segments that he may not now service.

This paper describes a methodology for retail fashion market monitoring which applies fashion segmentation research techniques and retail image methods to the men's fashion context.

This methodology provides new input for the retailer in terms of identifying and profiling fashion market segment, in selecting his product lines, his merchandising approach, and in designing and targeting his entire retail presentation to specific fashion market segments.

Market monitoring permits the diagnosis of the market structure and the understanding of patterns of buying behavior over time and enables the retailer to precisely pinpoint his strengths and weaknesses as they currently exist and as they change over time.

- Anderson, R. C., and R. A. Scott. "Supermarkets: Are They Really Alike?" Journal of Retailing, 46: 16-24, Fall, 1970.
- Arnold, S. J., and D. J. Tigert. "Market Monitoring Through Attitudes Research", Journal of Retailing, Winter, 1973-1974.
- Arons, Leon. "Does TV Viewing Influence Store Image and Shopping Frequency?", Journal of Retailing, 37:1-13, Fall, 1961.
- Baumgarten, S. A. "The Innovative Communicator in the Diffusion Process." Journal of Marketing Research, 12:12-18, February, 1975.
- Berry, Leonard J. "The Components of Department Store Image: A Theoretical and Empirical Analysis." Journal of Retailing, 45:3-20, Spring, 1969.
- Grindereng, Margaret P. "Fashion Diffusion." Journal of Home Economics, 59:171-174, March, 1967.
- Kelly, R. F., and R. Stephenson, "Semantic Differential: An Information Source for Designing Retail Patronage Appeals," Journal of Marketing, Vol. 31, pp. 43-47, October, 1967.
- King, C. W. "Fashion Adoption: A Rebuttal to the 'Trickle Down' Theory," Toward Scientific Marketing, ed. S. A. Greyser. Proceedings of the Winter Conference of the American Marketing Association, Chicago: American Marketing Association, 1963.
- King, C. W., and J. O. Summers, "Overlap of Opinion Leadership Across Consumer Product Categories," Journal of Marketing Research, 7:43-50, February, 1970.
- King, C. W., D. J. Tigert, and L. J. Ring. "Contemporary Fashion Theory and Retail Shopping Behavior: A Segmentation Analysis." 1975 Combined Proceedings of the American Marketing Association, ed. E. M. Mazze, The American Marketing Association, Chicago, 1975.
- Lazer, W., and R. Wyckham. "Perceptual Segmentation of Department Store Markets." Journal of Retailing, Summer, 1969, 3-14.
- Martineau, Pierre. "The Personality of the Retail Store." Harvard Business Review, 36:47-55, January-February, 1958.
- May, E. G. "Psychographics in Department Store Imagery," Marketing Science Institute Working Paper, October, 1971.
- Painter, J. J., and K. L. Granzin, "Profiling the Male Fashion Innovator--Another Step." Advances in Consumer Research, Vol. III, ed. B. B. Anderson, The Association for Consumer Research, 1976.
- Pasnak, M. F. O. and Ruth W. Ayres. "Fashion Innovators," Journal of Home Economics, Vol. 61, No. 9, pp. 698-702, November, 1969.
- Reynolds, F. D., and W. R. Darden. "Fashion Theory and Pragmatics: The Case of the Midi," Journal of Retailing, 49:51-62, Spring, 1973.
- Ring, Lawrence J. "A Multiple Discriminant Analysis Approach to the Development of Retail Store Positioning." Colgate Darden Graduate School of Business Administration Working Paper No. 76-09, The University of Virginia, 1976.
- Ring, Lawrence J. "The Male Fashion Consumer: An Analysis of Fashion Involvement and Retail Patronage Determinants." Unpublished Doctoral Dissertation, Purdue University, West Lafayette, Indiana, 1977.
- Shapiro. B. "Price As An Indicator of Quality," Unpublished Doctoral Dissertation, Harvard Graduate School of Business, 1970.
- Sproles, G. B., and C. W. King. "The Consumer Fashion Change Agent: A Theoretical Conceptualization and Empirical Identification." Institute for Research in the Behavioral, Economic, and Management Sciences, Purdue University, Paper No. 433, November, 1973.
- Stephenson, P. R. "Identifying Determinants of Retail Patronage." Journal of Marketing, July, 1969, 57-61.
- Tigert, D. J., Terry Cotter, and Sylvia Ma. Mom Always Liked Ralph's Best: The Strategic Crisis for Safeway in California, Monograph #2. The Retailing and Institutional Research Programme, The University of Toronto, 1976.
- Tigert, D. J., and L. J. Ring, and C. W. King, "Fashion Adoption and Buying Behavior: A Methodological Study," Advances in Consumer Research, Vol. III, ed. B. B. Anderson, pp. 46-52, The Association for Consumer Research, 1976.
- Weale, B. W. "Measuring the Customer Image of a Department Store," Journal of Retailing, 37:40-48, Summer, 1971.
- Wyckham, R. G., "Aggregate Department Store Images: Social and Experimental Factors," Unpublished Doctoral Dissertation, Michigan State University, 1967.

George B. Sproles, University of Houston

Abstract

This paper presents a preliminary introduction to a survey research methodology for integrating some new techniques of fashion trend monitoring and retail store image/patronage monitoring into a retailing management information system. Selected preliminary findings from a two year longitudinal study of four separate random samples of consumers are used to illustrate the methodology.

Introduction

A new survey methodology for measuring retail store images and associated dimensions of store patronage has recently received increased use in retailing research (i.e. Arnold and Tigert, 1973; King, Tigert and Ring, 1975).

This methodology for market monitoring has been found to have a substantial degree of reliability and validity in describing the competitive positions of retail stores within a local market trading area. More recently the method has also proved useful for multivariate modeling and "mapping" of the structure of local retail markets (i.e. Ring and King, in press).

One use of this new methodology which has not yet been investigated involves integrating the methodology to some newly evolving techniques for measuring changing trends in consumers' product preferences and actual buying behavior. However, this represents the next logical step in the development of a comprehensive retail market monitoring system. Retailers need information not only on their competitive positions across general store image/patronage dimensions, but also on specific competitive products offered within their trading area. This type of data will ultimately be a key part of future retailing management information systems used by retailers as an aid in developing their general positions ("images") in the market and complementary product assortments.

Given these needs, the purpose of this paper is to describe a longitudinal survey research program which has merged the evolving market monitoring methodology with evolving techniques for measuring current and predicted future growth in consumer demand for specific products. The product category involved in this research is women's clothing fashions, which is one of the most difficult and yet crucial categories of products in which to measure consumers' current and predicted future preferences for consumption. This paper will present a brief overview on the development of the research methodology for the investigation. Some preliminary illustrative data on use of the integrated system will also be discussed.

¹Field administration of this study was funded by the Agricultural Experiment Station, Purdue University. Data analysis and computer programming was funded by the Department of Human Development and Consumer Sciences (pending), University of Houston. Mrs. Jacqueline McCullough was a doctoral researcher in one phase of this investigation.

Overview of the Methodology

The research was conducted in 1976 and 1977 in Lafayette, Indiana. In each year two separate random samples, one of college women at Purdue University and one of adult women in the Lafayette community, were mailed a questionnaire surveying their store patronage and fashion preferences. In the next paragraphs the store patronage measures, fashion preference measures, and administration of the study will be described.

The Store Image/Patronage Measures

The new retail image and patronage research methodology has been described in earlier cited references, and requires only brief introduction here. Basically, the methodology measures consumers' images of and preferences for retail store in their trading area across a variety of relevant patronage dimensions (i.e. store with the best location, best advertising, best assortment, most fashionable selections, store most often shopped, etc.). King, Tigert and Ring (1975) have developed a specific sequence of sixteen questions for use in clothing stores, and their sequence is used in the research reported in this paper.

A group of ten directly competitive clothing stores was selected for inclusion in the survey instruments, including 3 department stores, 6 specialty stores, and 1 chain store. The 10 stores had previously been identified in a pretest as among the most extensively patronized clothing stores in Lafayette. Each subject was asked to pick the best store on each image/patronage dimension from the list of 10 stores. Additionally, respondents were given a write-in for "other store," to measure instances when none of the 10 stores was perceived as best on a specific dimension.

The Fashion Preference Measures

The methodology for measuring consumers' fashion preferences and predicted buying behavior is relatively novel, and will be described in greater detail. The basic method was developed as a modification of techniques previously used by Baumgarten (1975). His method involved presenting subjects a series of slightly varied basic styles of dress (i.e. styles of pants, shoes), and asking each subject a series of consumption-related questions on each style. Those questions included number of items of each style currently owned, current popularity of each style, predicted future popularity of each style, and future buying intentions for each style. Selected measures were then statistically combined into measures of consumer innovativeness, and were used to identify fashion innovative consumers. However, it appears this type of methodology, with the appropriate modifications, can be used as a simple system of market monitoring and prediction of forthcoming fashion trends.

Since the technique for monitoring consumers' fashion preferences used in the present study represents a substantially different application of Baumgarten's approach, the development of this application will be described in some detail. The principal difference is that the technique is used to comparatively measure the current and changing fashionability of specific styles. Therefore, it is first necessary to validate the use of such a technique in discriminating between styles which are at different levels of current and predicted future fashion-

ability. A second major difference is that simple descriptive statistics (frequencies) are used to differentiate the fashionability of one style versus another. Thus the interpretation of findings is relatively uncomplicated and responsive to managerial needs.

To initially validate the use of this methodology in measuring stages and levels of fashionability for specific styles, a series of pilot pretests was conducted in the late Winter-early Spring of 1976.² In the first test, female college students were presented illustrations of over 150 styles of dress taken from various fashion periodicals, and were asked to classify each illustration into four categories.³ The categories can be described as follows:

- #1 -- The "newest" general fashion trend
- #2 -- Out of fashion
- #3 -- The "newest" college age/youth-oriented fashion
- #4 -- "Classic" (long established) fashion

For 42 styles which were placed in a single category by 50 percent or more of the first test subjects, a second test was run. This test used a group of junior-senior female students majoring in fashion retailing, who were used as "expert judges" to validate the categorizations done by the first test group. Based on the results given by this "expert" group, four dress styles and four pant styles representing the four categories were selected for use in the survey.

The selected styles were redrawn to the same size and scale on a neutral posed form, to control for any differences in appearances which could be attributed to anything other than the style itself (i.e. a model's pose, hair style, figure, facial expression). The redrawn styles were then subjected to a final validating test with female students. This test successfully validated each of the four drawings of dress and four drawings of pant outfits as fitting in one of the four categories of fashionability, though this test did indicate that minor redrawing of a few details on two outfits was needed.

To construct the questionnaire for this part of the survey, the four dress styles were randomly arranged in the left column of the page. The right column contained the four question sequence on consumption and perceptions of popularity for each style. The pant styles were then arranged in an identical manner on a separate page of the questionnaire.

Questionnaires containing the fashion preference and store patronage measures, as well as a variety of other measures not relevant to the present study, were mailed to separate random samples of 400 female students at Purdue, and 400 adults in Lafayette. The mailings were purposefully administered in the early part of the Spring 1976 fashion season, to obtain an "early season tracking" of fashion preferences. Additionally, the mailing was made to follow immediately after the pilot validation of categories for the four styles, thus avoiding the possibility that an abrupt change of fashion would invalidate any of the four categories.

²McCullough was a co-researcher and doctoral student who conducted the pretests. Also see McCullough (1977) for an analysis of fashion-conscious college students identified in the 1976 phase of this study.

³The concept of categorizing styles in a similar manner has previously been used by Reed (1973) in a study of clothing as a symbolic indicator of the self-concept of college women.

In the early Spring of 1977, questionnaires containing identical measures of fashion preferences and store patronage were mailed to random samples of 400 female students at Purdue, and 400 adults in Lafayette. Thus the survey administrations in 1976 and 1977 were directly comparable in terms of samples, questionnaire designs, and times of administration within the Spring fashion seasons of both years.

Other methodological details were as follows. In both the 1976 and 1977 surveys, the college student subjects included in the surveys were randomly selected from the Purdue registrar's list of full-time undergraduate female students. The adult subjects were randomly selected from the Lafayette telephone book. Representatives of the telephone company indicated that probably less than 5 percent of the Lafayette area households had unlisted phones or no phones; therefore, it would appear that the Lafayette telephone book was a relatively representative sampling frame for the purposes of this research. Finally, one methodological difference between the 1976 and 1977 studies should be noted. In 1976 the surveys were followed 10 days later by a reminder postcard which was used to stimulate increased response rates; however, the 1977 study involved a single mailing of a shortened questionnaire without use of follow-up procedures. This resulted in a difference in response rates in the two years of the study. The final response rates, in terms of returned usable questionnaires, was as follows:

1976 - College students	250 returns (62.5%)
1976 - Adults	174 returns (43.5%)
1977 - College students	198 returns (49.5%)
1977 - Adults	122 returns (30.5%)

From these results it would appear that the postcard follow-up yielded an increased response in the order of 10 percent or more. This was found to be the case in analyzing the returns in the 1976 study by date of return as compared to the postcard mailing date.

Findings

The following discussion focuses on a preliminary analysis of data on consumers' preferences for the four selected dress styles, and patterns of preferences across store patronage dimensions. These data illustrate the validity of this new methodology for measuring fashion preferences and identifying the predicted buying behavior of consumers frequently patronizing specific stores.

Tables 1 - 4 present descriptive data for the fashion preference measures, in terms of raw percentages of subjects in each sample who responded to each consumption-oriented question. The numbers #1 through #4 in these tables are keyed to the categories of fashionability described earlier in this paper. In general, the findings in these tables would significantly appear to validate the use of this monitoring technique in discriminating current levels of fashionability and perhaps in forecasting the forthcoming fashion trends. However, the latter point will require further research for more formal evidence on that point. Another point of interest is the divergence in preferences between the adult and college subjects. There are substantial differences in orientations toward fashion consumption of these two broadly defined market segments, and the surveys appear to discriminate these differences. This is especially notable in the measures of preference for style #3, which was categorized as a college or youth-oriented fashion.

Some of the more notable findings in Tables 1 - 4 are as follows. In the Table 1 data on ownership of the four styles, style #4 (the classic fashion) would appear to be the most widely owned style, particularly among adults. College students were significantly less likely than adults to report owning #2 (out of fashion). Finally,

TABLE 1
Ownership of Selected Dress Styles Among Adult and College Women,
1976 and 1977 Surveys

	Number Owned, By Styles	Adult Women		College Students	
		1976	1977	1976	1977
#1	None *	53.4%*	41.8%	45.6%	38.4%
	1 or 2 *	36.2%*	43.4%	42.0%	51.5%
	3 or more*	10.3%*	14.7%	12.4%	10.1%
#2	None	55.7%	56.6%	73.6%	77.8%
	1 or 2	28.7%	30.4%	22.8%	17.7%
	3 or more	15.4%	13.1%	3.6%	4.5%
#3	None	66.7%	67.2%	54.0%	46.5%
	1 or 2	24.7%	22.1%	32.4%	42.0%
	3 or more	8.6%	10.6%	13.6%	11.6%
#4	None	36.8%	23.0%	39.2%	39.4%
	1 or 2	42.6%	50.0%	46.4%	45.5%
	3 or more	20.7%	27.0%	14.4%	15.1%

*Read: 53.4% of adult women in the 1976 survey indicated they owned no items of style #1, 36.2% indicated they owned 1 or 2 items of style #1, and 10.3% indicated they owned 3 or more items of style #1. Percentages in each such group of percentages in Table 1, and in Tables 2 - 4, do not add to 100% due to missing responses and/or rounding errors.

TABLE 2
Current Popularity of Selected Dress Styles Among Adult and College Women,
1976 and 1977 Surveys

	Degree of Popularity, By Styles	Adult Women		College Students	
		1976	1977	1976	1977
#1	Extreme	45.4%	41.0%	72.8%	44.9%
	Moderate	44.3%	51.6%	25.2%	53.5%
	Not	8.6%	4.1%	2.0%	1.0%
#2	Extreme	9.2%	4.9%	1.6%	2.0%
	Moderate	44.3%	34.4%	25.6%	25.3%
	Not	43.7%	56.6%	72.8%	72.2%
#3	Extreme	20.1%	16.4%	41.6%	27.8%
	Moderate	47.1%	50.0%	41.2%	54.5%
	Not	29.9%	28.7%	17.2%	17.7%
#4	Extreme	35.6%	50.0%	18.4%	44.9%
	Moderate	48.3%	39.3%	57.6%	49.5%
	Not	13.8%	6.6%	24.0%	5.6%

TABLE 3
 Expected Popularity of Selected Dress Styles Among Adult and College Women,
 1976 and 1977 Surveys

	Degree of Expected Popularity, Next 6 Months, By Styles	Adult Women		College Students	
		1976	1977	1976	1977
		#1	Less	12.6%	13.9%
	Same	59.8%	52.5%	47.6%	63.6%
	More	23.0%	27.0%	42.8%	20.7%
#2	Less	33.9%	39.3%	38.4%	56.1%
	Same	55.7%	45.9%	53.2%	36.9%
	More	5.2%	7.4%	6.4%	4.5%
#3	Less	33.3%	32.8%	15.6%	23.2%
	Same	43.7%	41.8%	50.0%	49.5%
	More	17.8%	20.5%	32.4%	26.3%
#4	Less	12.6%	10.7%	23.2%	16.2%
	Same	55.2%	58.2%	58.8%	60.6%
	More	27.6%	26.2%	15.6%	21.7%

TABLE 4
 Intended Purchases of Selected Dress Styles Among Adult and College Women,
 1976 and 1977 Surveys

	Intent to Purchase, By Styles	Adult Women		College Students	
		1976	1977	1976	1977
		#1	Definitely	10.9%	12.3%
	Possibly	35.1%	42.6%	52.8%	56.1%
	Will Not	51.7%	41.8%	24.8%	30.3%
#2	Definitely	3.4%	2.5%	2.4%	2.0%
	Possibly	22.7%	19.7%	9.2%	7.6%
	Will Not	70.7%	74.6%	86.4%	88.4%
#3	Definitely	4.0%	7.4%	12.0%	8.6%
	Possibly	16.7%	20.5%	40.0%	40.4%
	Will Not	77.6%	67.2%	47.2%	50.0%
#4	Definitely	9.2%	11.5%	8.8%	20.2%
	Possibly	35.1%	44.3%	33.2%	41.9%
	Will Not	52.3%	40.2%	56.8%	36.9%

growth in ownership of #1 (the newest general fashion) can be seen from the 1976 to 1977 studies for both the adult and student samples. Thus these ownership measures give a distinct picture of consumers' preferences in the market (this of course becomes even more true upon visual inspection of the line drawings for each style).

Tables 2 and 3 show consumers' judgments of current and predicted future popularity of each style. In the 1976 data of Table 2, it is clear that both the adult and college samples rated the category of fashionability of each style in the same general categories as identified in the pilot tests. This represents a significant validation of the methodology for discriminating current fashionability of styles based on using basic line drawings of the styles. In the 1977 data, results are similar, though there is a clear decline in popularity ratings for #1 and #3 (which were highly popular in 1976). The Table 3 data showing predicted popularity for each style in the next six months is also of considerable interest. Among adults the prediction is for #1 (the newest general fashion) and #4 (classic) to grow most significantly in popularity. In comparison, students clearly pick #1 and #3 (the youthful fashion) for the most significant growth. Finally, both adult and college groups indicate that #2 (out of fashion) is leveling out or declining in fashionability.

Table 4 presents data on consumers' purchasing intentions for each style. Although the preceding tables indicated many consumers would seem to have relatively well formed perceptions regarding popularity of each style, there is uncertainty as to what will actually be purchased. This is indicated by the high percentage of consumers responding they would "possibly" purchase each style in the near future. Only for style #2 (out of fashion) is the trend clear, and in this instance approximately 75 percent or more of consumers indicated they would not purchase the style. The styles most likely to be purchased by adults appear to be #1 and #4, while college students appear to be most positive about #1 and #3. It is also interesting to note the significant growth in students' purchase intentions for #4 (classic fashion) from the 1976 to 1977 study, as well as their apparent decline in intentions regarding #1 (newest general fashion in 1976).

Tables 5 and 6 exemplify how data from the fashion preference monitoring system described in Tables 1 - 4 can be integrated with data from retail store image/patronage measures. Tables 5 and 6 contain cross-tabulations of three stores which were identified by subjects as "most frequently patronized" versus the style preferences expressed by consumers choosing each store. The three stores selected for this example illustration include a well-known local department store (store A), a fashionable local specialty store (store B), and a national chain store (store C). These three stores were found to be among the most frequently patronized stores in the Lafayette area, and they collectively controlled a major share of the total clothing market. The three stores also represent different organizational structures, marketing strategies, and store images in the local market.

The percentages in Tables 5 and 6 represent the proportion of consumers choosing each store as most frequently patronized who rated each style as "high in current popularity," "will increase in popularity in next six months," and "definitely will purchase" this style. These percentages, when compared across styles, give an indication of the specific styles which would appear to be highest or growing in popularity, and which therefore should be emphasized in each store's inventory. The data also give an approximation of the proportion of inventory which might be committed to each style. The general finding of these tables is that consumers who have a "favorite" (most frequently shopped) store can

have different fashion preferences compared to those choosing other stores. A brief description of these differences in preferences, and their implications for the buying and assortment policy of each store, is as follows:

- 1) Store A - Department Store. The 1976 data indicate this store has substantial appeal to both adult and college markets. In adult-oriented departments, the data clearly indicate that emphasis should be placed on styles #1 and #4, perhaps including a balanced inventory of those styles. In comparison, the data indicate that college-oriented departments should carry styles #1, #3 and #4, with some weighting toward #3. The 1977 data indicate a general decline in styles #1 and #3, and a considerable growth in importance of #4, for both the adult and college segments of the market.
- 2) Store B - Specialty Store. The 1976 and 1977 data indicate this store appeals largely to the college market segment, with approximately 3/4 of the regular patrons coming from that market segment. The 1976 data indicate the store's inventory should be heavily weighted toward style #1, though a modest amount of #3 and #4 should be carried in stock. It is apparent that the emphasis on #1 would appeal not only to the college market, but also to the few adults favoring this store. In the 1977 findings there is a substantial shift in consumers' preferences, and it would appear that a relatively balanced inventory including styles #1, #3 and #4 would be best, though a slight over-representation of #3 and #4 would be acceptable.
- 3) Store C - Chain Store. In the 1976 data there are indications that this store has a substantial competitive appeal in both adult and college markets, and the market share of this store would appear to compare favorably to stores A and B. The data also indicate that this store should carry the widest variety of styles in order to appeal to its clientele. There are some clear indications that the most emphasis should be placed on style #1 (the newest fashion), but even style #2 (the out of fashion style) appears to achieve enough popularity and buying intentions to justify some modest representation in the store's inventory. The 1977 data also appear to support this policy of variety and balance for the chain store. It would thus appear that this chain store has some of the most complex inventory decisions to make, given its widespread appeal to the consumer market and the variety of styles appealing to those market segments. Considerable coordination between local store management and the centralized management of the chain would clearly be needed.

Discussion

This research has validated a mail survey technique for monitoring trends in consumers' fashion preferences. The technique involves presenting consumers with line drawings of basic styles, and obtaining consumers' rating of their ownership of each style, popularity (current and future) of each style, and future purchasing intentions. The data indicate that responses to these types of questions can be used to discriminate the current fashionability of each style. Predictions of future demand can also be extrapolated from the data, though the validity of those predictions has not yet been fully explored.

TABLE 5
Style Preferences Among Adult Women, 1976 and 1977 Surveys,
According to Store Most Often Shopped

Popularity Ratings, Purchase Intentions			1976 Stores*			1977 Stores*		
			A	B	C	A	B	C
Current Popularity (High)	Styles	#1	47% ^a	50% ^a	61% ^a	29%	29%	42%
		#2	6	8	13	3	0	0
		#3	28	17	22	15	29	8
		#4	38	17	26	62	57	50
Predicted Popularity (Will Increase)	Styles	#1	20	42	13	31	29	19
		#2	6	0	13	3	14	12
		#3	20	8	13	21	43	8
		#4	33	17	13	27	43	31
Purchase Intentions (Definitely Will Purchase)	Styles	#1	11	25	13	9	29	0
		#2	0	0	4	0	14	8
		#3	6	0	4	0	14	0
		#4	15	8	4	6	43	11
Base Number of Consumers			(49)	(12)	(23)	(34)	(7)	(26)

* A = Major Local Department Store
B = Specialty Store
C = National Chain Store

^aRead: In the 1976 study 47% of adult customers choosing store A as "most often shopped" rated current popularity of style #1 as high. Similarly, 50% of store B customers rated #1 high, and 61% of store C customers rated #1 high.

TABLE 6
Style Preferences Among College Women, 1976 and 1977 Surveys,
According to Store Most Often Shopped

Popularity Ratings, Purchase Intentions			1976 Stores*			1977 Stores*		
			A	B	C	A	B	C
Current Popularity (High)	Styles	#1	79%	73%	81%	32%	48%	48%
		#2	0	0	0	0	0	3
		#3	49	51	38	26	37	23
		#4	12	14	28	45	44	65
Predicted Popularity (Will Increase)	Styles	#1	27	53	74	21	15	27
		#2	6	11	3	5	7	3
		#3	22	44	32	24	30	20
		#4	22	11	13	30	19	30
Purchase Intentions (Definitely Will Purchase)	Styles	#1	6	32	28	3	11	23
		#2	3	6	0	0	4	3
		#3	18	14	6	8	19	10
		#4	13	8	6	19	15	29
Base Number of Consumers			(33)	(37)	(32)	(38)	(27)	(31)

*See interpretive notes at the bottom of Table 5.

The fashion preference monitoring technique can also be used to identify styles preferred by consumers who patronize specific stores most frequently. This type of information, when available to the retailer with the proper lead time, can be extremely useful as a merchandise management information system. Specifically, the data can be used to identify basic styles to buy, proportions of inventory to commit in each style, market segments for each style, and styles nearing the end of their period of fashionability. These types of information are crucial for effective and timely inventory management.

The fashion preference monitoring technique can be especially useful for longitudinal tracking of trends. In the present research a two year tracking was conducted, with studies conducted on an annual basis. However, to more precisely track changing trends, it is likely that more frequent studies would be necessary to fully respond to the needs of retailers. For example, a monitoring system keyed to major annual fashion seasons (Spring and Fall) might be more responsive to retailer needs. Furthermore, within any given season it would be useful to obtain an initial measurement of trends at the start of the season (as done in the present study), and then conduct a follow-up study at mid-season in order to monitor the trend and validate the earlier findings. Such accumulations of data, when compared to the store's actual results, could be very useful in precisely validating the predictive efficacy of the monitoring system for any given store.

Finally, it is important to note that the data presented in this paper are only one example illustration of how fashion preference and store patronage monitoring systems can be integrated into a management information system. In this example the focus was on the store patronage dimension of "store most frequently shopped." Similar comparisons such as those made in Tables 5 and 6 could also be made across other relevant store image/patronage dimensions (i.e. store last shopped, most fashionable store store with best assortment, store with best prices, store with the best location, and so on). In this manner a complete profile of consumers' fashion preferences and their view of the competitive market can be obtained. This type of competitive information can be crucial for retailers who wish to assess their position in the local market, and it can be absolutely vital to those retailers who need to reposition their images in order to maintain their share of the local market.

References

Stephen J. Arnold and Douglas J. Tigert, "Market Monitoring Through Attitude Research," Journal of Retailing, 49 (Winter, 1973-1974), 3-22.

Steven A. Baumgarten, "The Innovative Communicator in the Diffusion Process," Journal of Marketing Research, 12 (February, 1975), 12-18.

Charles W. King, Douglas J. Tigert and Lawrence J. Ring, "Contemporary Fashion Theory and Retail Shopping Behavior: A Segmentation Analysis," Paper presented at the Marketing Educators' Conference of the American Marketing Association, Rochester, New York, August 17-20, 1975.

Jacquelyn Harden McCullough, A Multivariate Profile of Fashion Conscious College Women, Unpublished Ph.D. Dissertation, Purdue University, 1977.

Julia Ann Reed, Clothing As A Symbolic Indicator of The Self, Unpublished Ph.D. Dissertation, Purdue University, 1973.

Lawrence J. Ring and Charles W. King, "A Multiple Discriminant Analysis Approach to the Development of Retail

Store Positioning," Advances in Consumer Research, 5 (1977), in press.

SEGMENTING LOCAL MARKETS FOR
ENTERTAINMENT SERVICES:
THE CASE OF DISCOTHEQUES

Daniel A. Emenheiser, Oklahoma State University
George B. Sproles, University of Houston

Abstract

There are many and varied entertainment services available today. Retailers of entertainment face problems and challenges concerning local market segmentation, focusing on recognizing and then satisfying the needs of their particular consumers. This paper presents an analysis of these problems and illustrates a method of segmenting local markets for such entertainment services, focusing on the segments patronizing two competitive discotheques in Indianapolis, Indiana. An effective profile of the "heavy user" market segments of these two discotheques is developed as a result of the research findings. There are some significant similarities between these segments, but there are also some very important differences between "heavy users" of the two discotheques, which lead to market segmentation implications.

Introduction

One aspect of retailing which receives little attention is the marketing of entertainment services. Nevertheless, the providers and promoters of entertainment services such as night clubs, discotheques, resorts, theme parks, health spas, tennis clubs and so on are very much in the business of retailing. They also face the retailer's classic problems of identifying local markets of consumers, analyzing consumer preferences, responding to changing services of their competitors, and developing retailing strategy for their local markets and market segments.

This paper presents an analysis of the marketing problems faced by retailers in the entertainment business, focusing on the segmentation of local markets for entertainment. An approach to segmenting local markets for a specific type of entertainment is described, using a combination of established and new techniques for market segmentation research. The specific example used is discotheques, a relatively new form of entertainment which has recently grown to become an important part of the contemporary market for entertainment. The questionnaires utilized measured variables including discotheque attendance, discotheque preferences, general discotheque interests, factors important when choosing a favorite discotheque, activities, general interests, and the respondent's personal situation.

The purpose of this research was to construct a comprehensive profile of patronage characteristics of the customers who frequent two selected discotheques in the Indianapolis, Indiana area. Emphasis is placed on constructing a complete profile of the "heavy user" of the selected discotheques, including demographic, life-style, and patronage preferences. This type of information on the market segments patronizing a particular discotheque is fundamental to the formation of discotheque retailing strategy.

Background

Within the past two years, the appeal of discotheques has been growing rapidly. This is evidenced by the ex-

treme success of the numerous discotheques that have come into existence. There are an estimated 10,000 discotheques in the United States today, and last year discos claimed a \$3.6 billion share of the entertainment market (O'Harro, 1976). Major leaders in the hospitality industry have been developing the discotheque concept and incorporating it into their establishments. For example, innovative disco marketing has been pioneered by Hilton Hotels, Ramada Inn, Interstate United (the Giraffe discotheque chain), Bobby McGee's, and Marriott. This growing industry also includes new chains and innumerable single proprietorships in large and small cities throughout the United States.

Zucaro (1976) has noted that the restaurant/discotheque has become a multi-function entertainment center which is capable of creating various atmospheres and handling many different types of events. Discos, like other forms of entertainment, must cater to their following, and a major consideration is the type of clientele sought. As described by Orth (1976) a contemporary disco is a place for people to relax and "let go," to be seen, and to watch; it is also an environment where technology allows an anonymity over insecurities and hangups. However, this perseverance goes far beyond the actual discotheque. It has created a popular emphasis on fashion, dancing, and many other areas of modern life. Discotheques have also profoundly influenced the tremendous popular-music industry by creating a dynamic disco sound (i.e., Motown's primary concentration right now is on "disco" music), a vast market for records and entertainers, and a new all-disco format for radio stations.

There are many reasons why these restaurant and hospitality operations are developing the disco concept. However, the primary reason is because of the high profit returns possible. Many lounges that previously relied on live entertainment are now being turned into profitable disco centers, thereby encouraging customers to remain in the units well after dinner hours. With the rising cost and uncertain availability of top name live entertainment, this appears to be a viable way to maintain continued business. Many discos operate adjoining restaurants, and some units are also experimenting with the new trend of offering a special limited late night breakfast menu when the disco closes.

Discos frequently offer a variety of attractions to their clientele. The "typical" disco features dancing to popular music, played from professional sound systems operated by a disc jockey. The disc jockey is usually situated near the dance floor. Often this person interacts with the patrons through remarks, requests, and various contests. Music selections are often mixed to appeal to all age groups. The total "product mix" of the disco will frequently include a combination of non-stop music, professional stereo sound system, a dance floor, pulsating lights, phenomenal decor, plants, liquor, food, and much more. Backgammon tables, huge TV screens, slide shows, pinball and other amusement machines may also dot the disco "dancescape."

Because of their emphasis on phenomenal decor and design by many discotheque operations, a considerable number of prominent interior designers and architects have become involved in this industry. Many discotheques throughout the country have received national and professional recognition for their design and decor. This combination facilitates and helps to promote the image of the discos, which is thought to be of tremendous importance in drawing a clientele. For instance, Funky's Discotheque of Evansville, Indiana, was once the Holtz Steamboat Foundry built back in the 1860's. The decor of Funky's today exhibits the skilled craftsmanship of the 19th century. Every article from the stained and cut glass to the ornate wood workings are original antiques. Funky's is in constant search of antiques of this sort, and they also offer some of their antiques for sale to their customers as their menu indicates. Another example is the Library, a tri-level discotheque in Atlanta, Georgia, which boasts a half-million dollar interior. The walls of the establishment are bookshelves lined with books, which customers are actually allowed to check out.

In addition to the emphasized phenomenal decor and dancing, a central and highly profitable attraction for many discos is their adjoining restaurant. Many of these operations are receiving recognition for their excellence in food quality and selection. Unique, interesting foods are being served as a part of the overall disco entertainment package. This gives the disco management further flexibility in developing the products and services to be offered to the various market segments. Given the growing significance of eating away from home, this dimension of the disco's "product mix" can in fact become a major key for long-term financial success of the disco concept.

After interviewing prominent discotheque owners and managers, Tucker (1975) concluded that the secret to success for these professionals is to find something unique, something which will promote the "image" of their clubs. This virtually dictates the disco's emphasis on costly interiors, specialized decor, multi-level operations, high-powered sound systems, and special lighting. Some club owners are going to great expense to find the magic formula to satisfactorily meet the demands of the dominating affluent, young-executive market segment, but at the same time to have a disco that will bridge all age groups. This is possible because the versatility of discotheques is unique resolves Jarigue (1975). There are many possibilities available to management in developing the product/service mix for the discotheque. By developing a unique product/service mix, customers may be satisfied. But first management must know who comprises their market segments, the characteristics of these people, and what they really want from the disco. When management acquires this information, then each disco can develop its own formula to approach the desired market segment(s). Profitability of today's discos results in part from the special promotion and merchandising offerings and campaigns of each successful discotheque.

Delineated in the broadest of terms, the principal clientele of discotheques comprise what is sometimes referred to as the "Hungry Generation." The "Hungry Generation" represents the post-World War II bumper baby crop that is now growing into young adulthood, family life, and career employment. The largest part of this potential market appears to be young adults in their early to mid-twenties. However, a significant and perhaps growing segment is among established young adults ranging in age into the mid-thirties.

Nation's Restaurant News in an article titled "Feeding the Hungry Generation" (1971), has identified some specific attributes of the "Hungry Generation." These attributes include openness, involvement, spontaneity, and uninstitutionalism. Following through, today's "Hungry

Generation" is demanding choice and participation. They are looser and brutally honest. Other key orientations include impatience with tradition and refusal to be regimented. These characteristics suggest the hospitality and entertainment industry can increase its success in handling this young market by being less rigid and tradition-bound, and more flexible and innovative, in designing its product/service mix.

Discotheque and restaurant management must satisfy the desires stemming from these characteristics. Examples of some successful approaches include the implementation of salad bars, sundae bars, dance contests and lessons, contemporary decor, etc. Established chains have also adopted new restaurant images directed to the younger generation. A recent example of this occurrence is Howard Johnson's new restaurant system known as Ground Round. One would never know that a Ground Round is a converted "orange-roofer," but it is so successful (\$10,000 weekly sales are common) that the company is now building them from scratch. Although a Ground Round is principally a restaurant and bar, it has drawn upon the needs and desires of the "Hungry Generation" and has developed a product/service presentation which directly utilizes the characteristics of the discotheque concept. As Howard Johnson has done, many other established operations may continue their growth by finding fresh, new ways of expressing the new requirements that the young-thinking market segments are currently demanding.

Though discos have clearly achieved initial market acceptance among a rather young clientele, their market potential can be much broader. For example, Mr. Richard Chapman (1976), Assistant Director of Food and Beverage Planning and Development for Hilton International, remarks that, "The discotheque concept is a product, for which today's society in many locations has created a need. I feel that it is a product where the final mix must be finely attuned to the local market conditions. In general there is a trend away from the loud "rock" music to a more sophisticated and intimate level of operation. The former is more susceptible to the whims of fashion and appeals to the younger market, where spending powers are limited. The latter, perhaps requiring greater initial investment, appeals to the middle management mid-thirties age group." This suggests the importance of broadening the base of the disco's clientele whenever local market conditions permit.

Some notable attempts to broaden their markets have been made by various restaurant/discotheque operations. A principal emphasis is on diminishing the "youthful" image of discos, and making the modern disco a place where a variety of age/interest groups feel comfortable. Walton (1976), in a Nation's Restaurant News article titled "Hungry Generation Grows Up -- New Themes Aim at Wider Age Groups," points out several examples of this trend. For example, Bobby McGee's Conglomerate of discotheques considers customers aged 25 and up its basic target group and sees the concept's appeal as broad. Similarly, Steak and Ale views its market as "very broad," according to one company officer, and it has appealed to a 25-59 age group. Steak and Ale's Market Researcher, Albert J. Yesk, also identifies market segments within the population by values they hold rather than by age groups. Those expected to contribute to restaurant industry growth are "new values" people -- those whose interests and activities are not centered on home life. People, no matter what age, come to the discos to have fun, "to get away from it all" and to forget the hassles of everyday life. They do not come to get drunk, and they are mostly positive people as far as their entertainment and their social life go.

Disco operators who can fully understand the special identity and needs of the "Hungry Generation" will probably find themselves making extensive and more profitable uses of their capital investments. However, opera-

tors who can tap this "Hungry Generation" market will probably find it broader based than it first appears. Though the "Hungry Generation" may appear to be one homogeneously young "swinging" market, this is probably not the case. Rather, the disco market may contain many sub-segments, each having a relatively unique profile of personal characteristics and patronage preferences. In such a situation, disco marketing may be "tailor-made" to the special needs and wants of profitable sub-segments within this broad market.

Research Objective

Based on the preceding rationale, the basic objective of this research is to profile the characteristics of patrons of two major discotheques with restaurants in Indianapolis. The focus is on identifying the demographic, life-style, and patronage characteristics of "heavy users" of each discotheque. From these profiles it can be determined the extent to which the two clienteles of the discotheques are in fact different, and the extent to which the local market for discos can be segmented based on those differences. Finally, based on these findings managerial recommendations for planning a discotheque retailing strategy focused on the local market segments can be developed as an aid to managerial-decision making.

Methodology

Data were collected through a questionnaire which utilized dependent variables including, "About how frequently do you go to discotheques in the Indianapolis area?"; "Which disco do you patronize most often?"; and "When spending an evening in a discotheque, how much is your typical check?" Both life-style and demographic variables were included in the questionnaire to obtain specific descriptive data of the population and to aid in the development of a life-style profile of the "heavy user" of the discotheques.

The questionnaires used in this study were developed for application to discotheques based on previously conducted research. Discotheque preferences were measured using techniques developed by Arnold and Tigert (1973), and more recently modified by Charles King (Purdue University), Lawrence Ring (University of Virginia) and Douglas Tigert (University of Toronto). Questions from their instruments, which were designed to measure various patronage/store image dimensions for food and clothing stores, were adapted for the case of discotheques. Other sections of the questionnaire were adapted from the fashion and retailing research program of George Sproles (University of Houston) (Sproles, 1977).

As a preliminary exploratory investigation, an in-depth pre-test was conducted at three disco-type operations in Indianapolis on Wednesday evening, June 23, 1976 and on Saturday evening, June 26, 1976. Twenty pre-test questionnaires were distributed at each discotheque on each evening. Thus a total of 120 questionnaires were distributed for the pre-test. The pre-test questionnaire was randomly distributed to both males and females, after 9:00 P.M. until closing at all three operations simultaneously. At each discotheque on each evening, 10 questionnaires were distributed to men and 10 questionnaires were distributed to women. When customers were approached with a questionnaire, they were briefly informed of the purpose of the study, and then were asked if they would complete the questionnaire at their convenience. If they declined, they were thanked and no more was said. The process of distribution was designed not to disturb the clientele of the operations, and was done with cooperation of the management. The questionnaires were distributed along with an explanatory cover letter and an addressed, self-stamped return envelope.

Within thirty days the following response rates were obtained:

Discotheque A:	12/40 or 30%
Discotheque B:	12/40 or 30%
Discotheque C:	6/40 or 15%.

Overall, for the pre-test there was a 25% response rate, with a total of 10 male responses and 20 female responses.

Relevant information gained as a result of the pre-test was used to design a final questionnaire for the actual test. First, general response from the subjects and managers indicated that the length of the initial questionnaire kept a number of subjects from completing it. The original questionnaire was a four-page booklet, with questions on eight sides surveying the areas mentioned previously. Because of the length problem, the information contained on the first two pages of the questionnaire was shortened by the deletion of several of the discotheques. Several of those discos (other than the pre-test sites) went out of business between the pre-test and final test. These deletions allowed increased white space on the questionnaire, and aided in forming a shorter survey. The third page of the questionnaire which was concerned with certain aspects of personal attitudes and identity, the one requiring the most time to complete, was also deleted. With these changes, the actual questionnaire used still surveyed the principal research issues but did not appear nearly as lengthy or complex.

The pre-test results also resulted in a decision to remove Discotheque C from the final study, for several reasons. As indicated previously the response rate from Discotheque C was very low. This appeared to result from the fact that many of this disco's patrons were not true customers of the disco. Rather, many were dinner patrons and the disco or lounge area was often just a gathering or waiting place prior to dining. Consequently many of this disco's patrons felt they were not qualified to complete the questionnaire, and therefore they did not do so. Another reason for deletion of Discotheque C from the study was that the management did not add a dance floor to their lounge as they had planned at the start of the study. Therefore, this disco operation was not a true discotheque in the sense of offering the complete mix of music, dancing, and beverage service.

The pre-test also indicated the response of males in answering and returning the questionnaire was less than satisfactory. It was also tentatively found that there appeared to be minimal significant difference between the responses given by males and those given by females. Therefore it was decided to survey only the female patrons. This approach appears to have face validity, for many disco owners as well as the researchers believe that women are the crucial sex to satisfy in a disco retailing strategy. Likewise, Michael O'Harro (1976), one of this nation's leading discotheque consultants, supports the important role of women in the marketing process for hospitality operations. He suggests that if the women are satisfied and truly enjoy a given disco operation, then the male clientele will also be there, simply because the women are at the discotheque and are having a good time there.

The two competitive discotheques selected for final study are extremely successful discotheque/restaurant combinations. They are located in one of the affluent, upper-middle class suburban neighborhoods of Indianapolis. They are both located near popular shopping complexes, and are about a mile apart from each other. This close proximity makes them potential direct competitors for the same local market. For the actual test, 1,239 questionnaires were distributed to all cooperating female patrons who patronized either Discotheque A or Discotheque B during four survey nights. The nights were Wednesday-Saturday, July 21, 1976-July 24, 1976. Ques-

tionnaires were distributed by the same method used during the pre-test, from 9:00 P.M. until closing. Eight weeks were allowed for the return of the questionnaires.

The principal objective of this research was to identify the "heavy user" of the competitive discotheques. The dependent variable question "About how frequently do you go to discotheques in the Indianapolis area?", was used to identify "heavy users." Those who responded with "at least once every two weeks" were termed the "heavy users" of the discotheques. Additionally, another dependent variable question, "When spending an evening in a discotheque, how much is your typical check?", was used to identify "heavy spenders," a group of obvious interest to disco owners. The "heavy spender" was defined as that person who, when spending an evening in a discotheque, had a typical check of at least \$4.00 per person. The third dependent variable focused on the question "Which disco do you patronize most often?"

The "heavy user" of each disco was profiled across four major variable sets: 1) disco preferences, 2) disco image dimensions, 3) life-style activities/attitudes, and 4) demographics. Each profile variable was cross-tabulated against the selected "heavy user" dependent variable, using the .05 level for the test of significance. The following discussion will focus on a verbal summary of principal significant findings. See Emenheiser (1977) for a complete presentation of the data tables.

Results

A total of 1,239 questionnaires was distributed for the actual testing. A total of 754 questionnaires was distributed at Discotheque A and 485 questionnaires were distributed at Discotheque B. The greater number at Discotheque A resulted from the fact that the capacity of Discotheque A is greater than that of Discotheque B, and patronage was greater. After allowing eight weeks as a return period, 355 questionnaires were returned from the 1,239 which had been distributed, for an overall response of 29%. From the 754 questionnaires which had been distributed at Discotheque A, 192 were completed and returned, for a 25% response rate. From the 485 questionnaires which had been distributed at Discotheque B, 162 were completed and returned, for a 33% response rate.

A preliminary analysis indicated that 88.7% of the survey's 355 respondents have patronized Discotheque A, and 83.9% of the respondents have patronized Discotheque B. This provides assurance that the respondents are capable of effectively completing the questionnaire concerning these two competitive discotheques. It also implies (incorrectly as shown later) the market segments of the two discos actually view the two discos as competitive substitutes to one another.

Other initial findings were that the "heavy users" of the discotheques concentrate their patronage at one of the discotheques studied, as would be expected. However, the relation between the "heavy spender" and the variable denoting the disco patronized most often is more significant than is the relation between the "heavy spender" and the "heavy user" variables. This means that although the "heavy user" either patronizes Discotheque A or B most often, the "heavy user" is not always the one spending the most money at the disco. Thus it would appear that "heavy spenders" and "heavy users" of discotheques, as defined in this research, are not identical market segments.

The profile data on patronage preferences for "heavy users" of each disco yields some interesting findings. For each disco, "heavy users" said their chosen disco (A or B) was "best" across characteristics listed in

Table 1. However, competition between the two discos was high, as further analysis indicated. While "heavy users" of Discotheque A naturally rated it highly across most patronage dimensions, Discotheque B received the second highest number of "best" mentions across most dimensions. Likewise, among those preferring Discotheque B, Discotheque A receives second highest number of mentions.

TABLE 1
CHARACTERISTICS OF PREFERRED DISCOTHEQUE

is the easiest one to get to
has the lowest prices
has the most courteous service personnel
has the most efficient service personnel
has the highest quality food and/or beverage
gives you the best value for your money
is the most current, up-to-date
has the most exciting atmosphere
has the best selection of music
has the most enjoyable disc jockey presentation
has the best lighting system
has the best sound system
has their favorite dance floor

However there are some important differences in relation to other patronage dimensions. The patrons of Discotheque B believe Discotheque A is the disco with the highest prices, while they believe very strongly that Discotheque B has the most variety in food and/or beverage selection. The patrons of Discotheque A are evenly divided on this last factor. The patrons of both discotheques tend to agree that Discotheque A is the disco with the most exciting decor and most romantic atmosphere.

There are also a number of significant life-style/psychographic characteristics differentiating patrons of the two discos. Most interesting is the fact the "heavy users" of Discotheque A visit the disco for relaxation and enjoyment, while those at Discotheque B visit the disco to meet with friends and for entertainment. This goes along with the fact that the "heavy users" at Discotheque B frequently join in informal gatherings with friends. Those of Discotheque A do also but to a lesser extent. Other findings were that the "heavy users" at Discotheque A participate more often in outdoor activities than do those at Discotheque B. The "heavy users" of Discotheque B were also found to report lower levels of attending religious services or doing gardening. Finally, many patrons of Discotheque A perceived that they did not carefully watch how they spent their money, while patrons of Discotheque B were evenly split on this issue.

There are also a wide range of similarities in life-style orientations and patronage preferences of the "heavy users." All tend to go swimming regularly in the summer, to eat lunch at restaurants, and to go shopping with friends. There is a consensus among "heavy users" that the hours of operation of the local discotheques in Indianapolis are satisfactory. They believe that discotheques should have late night (after midnight) sandwich or breakfast type menus, and that good tips should be given for good service. The "heavy users" agree that a visit from a disco artist at the discotheque would be exciting. However, they do not seem to prefer the disc jockey's music to that of a live band. Nor do they usually stop at another restaurant, after leaving the disco, for a nitecap. Particular emphasis is placed on the importance of the disco's atmosphere, cleanliness, and efficiency and friendliness of the service staff. Also very important to the "heavy user" is dance floor space and hours for

reduced drink prices. Little importance is placed on the availability of pinball or other electronic games, backgammon tables, or slide presentations. Concerning other consumption-oriented interests of the "heavy user," the data indicate they do not perceive themselves to be opinion leaders in choices of fashions. However, the data indicate they keep their wardrobes up to date with current fashions by usually having one or more outfits of the very latest style. Finally, the "heavy users" appeared to not necessarily be heavy consumers of the albums or tapes of their favorite singers.

Demographic differences between "heavy users" of the two discos were not substantial. However, the age comparison does appear to be very important. Discotheque A had a greater proportion of patrons from the 25 to 34 age group than does Discotheque B, while Discotheque B tends to have more patrons from the 18 to 24 age group. This factor is a strong indication of the transitional differences between the "heavy user" market segments of the two operations. Otherwise, demographically the "heavy users" of the two discos are much the same. They tend to live in suburban Indianapolis in large apartment complexes. In general they are single and are employed.

Discussion

On the basis of significant data collected several market segments of the two discotheques can be described, including "heavy users," "heavy spenders" and disco most frequently patronized. However, most important is the market segment composed of the "heavy users" because they are the ones who are for many operations most important to the success of that operation. Even though their typical check averages may not be quite as high as those of the "heavy spenders," they are important for many reasons. They represent repeat business, they represent a large market segment, and they are excellent "promoters" of a disco, i.e., they tell their friends about their favorite disco and often even meet them there. As a result they are definitely a very profitable market segment to the operations. Often this market segment can become so loyal to a given operation that many discos become known as having a certain segment of people usually there, i.e., "the regulars" (many discos now offer memberships to such groups).

A rather effective profile of the "heavy user" market segments of the two discos emphasized in this research has been developed. There are some significant similarities between these segments. It has been noted that those patronizing either Discotheque A or B most often perceive their chosen disco to be "best" in a similar set of products and services (see Table 1). Along other lines, because "heavy users" of both Discotheque A and B are all broadly a part of the "Hungry Generation," it should not be surprising to find many of their attitudes, interests, activities, needs, and desires to be very similar. Further, because the two operations are competitive in location and in products and services offered, most of the customers have patronized both discos occasionally, though they clearly prefer one over the other. In short, the principal market segments of both discos have a relatively clear set of preferences, and their identifying characteristics overlap to a substantial degree.

However, there are some very important differences between "heavy users" of the two discos, and these differences suggest a basis for segmenting the local market. Most important is the age difference between the two segments. There are also important differences in perceptions of prices and variety of food/beverage selections. Finally, the patrons of one disco appear to go to it for "relaxation," while patrons of the other attend to "meet friends." Clearly the preceding dimen-

sions constitute a fundamental basis for segmenting any local market for entertainment.

The findings of this investigation make it rather apparent that both discos have generated a relatively loyal clientele, and have successfully offered a product and service mix satisfying their principal market segments. However, the findings also support some possible new retailing strategies that could be added as a part of their strategies to remain "unique." Both discotheques may want to assess the feasibility and profitability of offering lunch and late night (after midnight) sandwich or breakfast type menus (why not utilize those available kitchen facilities?). Results of the study indicated a demand for these services by disco patrons, and offering expanded meal service could broaden the retailing strategy of each operation. More promotional efforts could also be placed on encouraging evening patrons to come for dinner as well as dancing. It would also appear that "tie-in" promotions with other retailers (i.e., clothing stores) might offer unique opportunities for the discos which would be favorably received by their younger clientele. Many other "tie-ins" oriented toward the life-style aspects might similarly appeal to the young. Finally, it appears that both discos have their principal appeal to younger market segments, and both might therefore increase profitability and stability for the future by promoting their product/service mix to a broader range of the total adult market.

One final note regarding research methodology should now be made. The present study has adapted new retailing research techniques and some established market segmentation techniques to the special situation of segmenting local markets for entertainment. Relatively little of this type of research has previously been done by the entertainment and hospitality industry, but the need for such research in the modern competitive and fast changing services market is clear. The present methodology used in this study exemplifies one successful approach which has generated managerially-relevant information on disco strengths/limitations, the identity of principal market segments patronizing competitive discos, and new competitive ideas for disco retailing strategy. Such "marketing intelligence" is crucial in opening new markets for discos and to continually satisfy consumers' needs.

Conclusion

The contemporary discotheque has become one of the strongest, rapidly growing, and most profitable segments of the hospitality industry. The typical disco offers a combined entertainment product/service mix including food, beverages, music and disc jockey personalities. The potential market for disco entertainment is large, but centers on a younger clientele ranging from late teens to mid-thirties in age. The type of research summarized in this paper can be an aid to disco owners/managers in segmenting their local markets, identifying patronage characteristics of contrasting segments, and formulating "disco retailing strategies" focused on the needs and demands of their segments of the market.

References

- Stephen J. Arnold and Douglas J. Tigert, "Market Monitoring Through Attitude Research," *Journal of Retailing*, (Winter, 1973), 3-22.
- Richard Chapman, Assistant Director of Food and Beverage Planning and Development, Hilton International, The Waldorf-Astoria, New York, New York, (personal letter to Daniel A. Emenheiser), (March 16, 1976).

Daniel A. Emenheiser, "Discotheques -- Consumer Behavior, Patronage Characteristics," unpublished M.S. thesis, Purdue University, 1977.

"Feeding the Hungry Generation," Nation's Restaurant News, (October 11, 1971), 11-12.

Ralph Jarigue interviewed by David Tucker, "Disco Saturday Night," In City, (October, 1975), 24-29.

Michael O'Harro, "Disco Owners Say Flexibility is Key to Fad Profits," The Washington Star, (August 8, 1976), 1.

Maureen Orth, "The Disco Whirl," Newsweek, (November 8, 1976), 94-98.

George B. Sproles, "Clothing Orientations of Adult Women in Indiana," Agricultural Experiment Station, Purdue University, 1977.

Douglas J. Tigert, "The Fast Food Franchise: Psychographic and Demographic Analysis," Journal of Retailing, (Spring, 1971), 81-90.

David Tucker, "Disco Saturday Night," In City, (October, 1975), 24-29.

Stephen Walton, "Hungry Generation Grows Up -- New Themes Aim at Wider Age Groups," Nation's Restaurant News, (January 5, 1976), 18.

Al Zucaro, "Disco-Restaurant or Restaurant-Disco?" Discotheekin, (August, 1976), 26.

LOCAL RETAIL MARKET MONITORING:
STRATEGIC IMPLICATIONS

Charles W. King, Purdue University

Abstract

Monitoring of the local retail market competitive structure for the retailer's particular product category and the retailer's position and store image within that competitive milieu is essential to the retailer's strategic planning. This paper discusses the retailer's strategic planning needs, the concept of local retail market monitoring as a routine research procedure and presents an example of pragmatic local retail market monitoring with a discussion of strategic implications.

Retailing: The Strategic Marketing Problem

The basic strategic marketing problem confronting the retailer centers around the classic issue of market positioning. The retailer's strategic objective is to develop an integrated marketing program encompassing a wide range of marketing elements to create a market position within his competitive environment.

To establish a market position, the retailer strives to develop store personality or image. The store image should be analytically defined and designed to appeal to and be compatible with the store's target customers. The store image is the product of a set of components or attributes that are important in determining customer patronage within a particular retailing competitive market.

Consumer perceptions of store images, however, are very time related. Store images can change significantly over time due to changing retailer strategies, new competitive entries or other exogenous market variables. The stability of store images over time can also vary sharply across product categories. Each week, for example, the price image of an aggressive supermarket is under scrutiny in its featured, end-of-week prices. Supermarket price image therefore, can be very volatile.

The fashion image of a major department store, by comparison, may change only modestly over several seasons. Even under heavy strategic focus, the fashion image of a retailer may be an extremely stable, difficult to change dimension of retail patronage.

The critical question is, how is the retailer's store image actually perceived by the target customers compared with the firm's competitors in the local market over time?

Local Retail Market Monitoring:
A Conceptual Overview

To address the issue of retailer market position and customer perceptions, the retailer needs an on-going local retail market monitoring program. The concept of local retail market monitoring is a direct extension of traditional store image research.

Historical Store Image Research: The State of the Art

Since the now classic Martineau (1958) conceptualization of the "store personality or image", store image research has become extremely topical among leading retailers and marketing academics. Over the past 19

years, great strides have been made in gaining management interest in the concept of store image as a key market variable and as a pragmatic strategic merchandising tool.

Over that period too, empirical store image research developed as well. Historical store image research, however, has been largely single sample survey and cross sectional in research design. The research has tended to be exploratory in topic content with too little historical linkage or futuristic continuity.

More importantly, across the store image research, there has been little standardization in field methodology or measurement instrumentation. Consequently, despite a relatively large literature, there has been only a limited amount of generalization across the field until recently.

From the historical store image research to date, however, has emerged an expanding research tradition. Analysis of that tradition indicates that:

- 1) A conceptual framework for understanding and researching store imagery has evolved;
- 2) A cadre of increasingly sophisticated empirical researchers in leading retailing organizations and in academia has entered the store image research arena;
- 3) The historical research has identified major "general" dimensions of retail patronage or store image such as merchandise selection or assortment, merchandise quality, merchandise pricing, locational convenience, merchandise styling(fashion), service(general) and sales-clerk service;
- 4) The research has indicated that the key determinants of retail patronage vary across product categories. Therefore, store image research, to be most effective to retailing management must be both "general" and tailored and focused to be product category "specific";
- 5) An operational field research design and measurement methodology for cross-study research comparability has evolved in the published literature.^{1, 2}

¹For representative and significant literature and industry trade practice reviews, see Wyckham, Lazar and Crissy (1971), Lindquist (1974-1975), May (1974-1975) and Ring (1977).

²Several leading research oriented retailers have made impressive progress in adapting this store image research tradition to their specific proprietary image monitoring needs.

The Concept of Local Retail Market Monitoring

Local retail market monitoring, as advocated here, is intended to produce a comprehensive analysis of the basic competitive retail market structure for a particular product category over time. The monitoring system must measure the retailer's historical and current market position and store image perceptions among its target customers.

Equally importantly, the monitoring system must also measure shifts in the relative positions of competitors in the local retailing arena as the market structure changes. The system must track both the number and market positioning of the major retailers and the changing strategies operative in the market over time.

In its ultimate form, the local retail market monitoring system would produce a comprehensive perceptual "image map" of a local retail market place. Based on that map, shifts in consumer perceptions and the resulting changes in retail market positioning could be tracked. The causes for the shifts could then be explored by more focused research.

More specifically, the local retail market monitoring methodology proposed here would expand and systematize the more traditional store image research. Within a particular product category, the local retail market monitoring research paradigm would:

- 1) Identify the major "general" and product "specific" determinants of retail patronage that should be monitored;
- 2) Develop and standardize an operational measurement battery designed to track consumer perceptions on those selected key dimensions;
- 3) Design an economically viable consumer survey research plan for soliciting consumer perceptions of retail store images;
- 4) Implement an on-going, systematic local retail market monitoring program using standardized field methodology and measurement instruments;
- 5) Based on the initial monitoring survey, prepare a comprehensive perceptual store image map--a based-line description--of the local retail market structure in the subject product category;
- 6) Routinely repeat the standardized market monitoring survey and analysis procedure to track changes in the perceptual store image map reflective of changing competitive conditions;
- 7) Based on changes in the perceptual store image map, direct more focused, in-depth research on specific problem areas and re-direct the retailing strategy accordingly;
- 8) Continue the routine market monitoring and retail market monitoring and retail strategy adaptation.

Local Retail Market Monitoring: A Pragmatic Example

The Toronto Male Fashion Research Program: On-Going Local Retail Market Monitoring

In 1974, the author launched exploratory research in the integration of traditional fashion adoption research with the developing area of store image and retail patronage analysis in the Toronto, Ontario, Canada fashion

retailing environment. Ultimately the 1975 Toronto Male Fashion Research Project was formalized and implemented.³

The 1975 research built on and integrated established research methodologies in the fashion segmentation, retail store image and lifestyle/AIO areas. The central thrust of the research focused on segmentation of the male fashion market ranging from the highly involved fashion change agent down to fashion disinterested/uninvolved consumer. Based on fashion interest/involvement segmentation measures, general and specific shopping/buying behavior was monitored across each of the major Toronto men's wear retailers.

A critical feature of the 1975 Toronto Male Fashion Research Project that had major impact on the strategic significance of the research was the author's commitment to retailer involvement. The major retailers who were to be subjects under study in the project were all solicited to contribute to the design and pragmatic analysis of the research results.

In terms of research funding too, the research effort was conceptualized as a joint academic/industry project. Purdue University and the University of Toronto contributed to the project through doctoral fellowships, University supplied computer time and professional support.

Industry funds, however, were solicited from local subject Toronto retailers to support the data collection and analysis costs. These costs were shared across all sponsoring firms. Therefore, to attract adequate field research funding from retailing management, the research design and analysis plan had to reflect strategic value to the sponsors at the outset of the research planning.

The basic field methodology involved two surveys of 1000 males each executed via a mail, self administered questionnaire based on area probability samples in Census Metropolitan Toronto. The two surveys were conducted in April - May, 1974 and April - May, 1975 to support longitudinal analysis of changes in store images over the 12 month period.

The concept of local retail market monitoring has been adopted by many of the major men's wear retailers in Toronto, largely as an out-growth of involvement in this research project. The 1975 project was launched as a one-survey extension of the 1974 pilot study. From that effort, a doctoral dissertation was produced (Ring, 1977).

At this point, however, the research effort has been syndicated as a commercial venture and expanded beyond its original academic objectives. The program is now entering its fifth year of operation and has proven the validity of local retail market monitoring as a pragmatic planning tool.

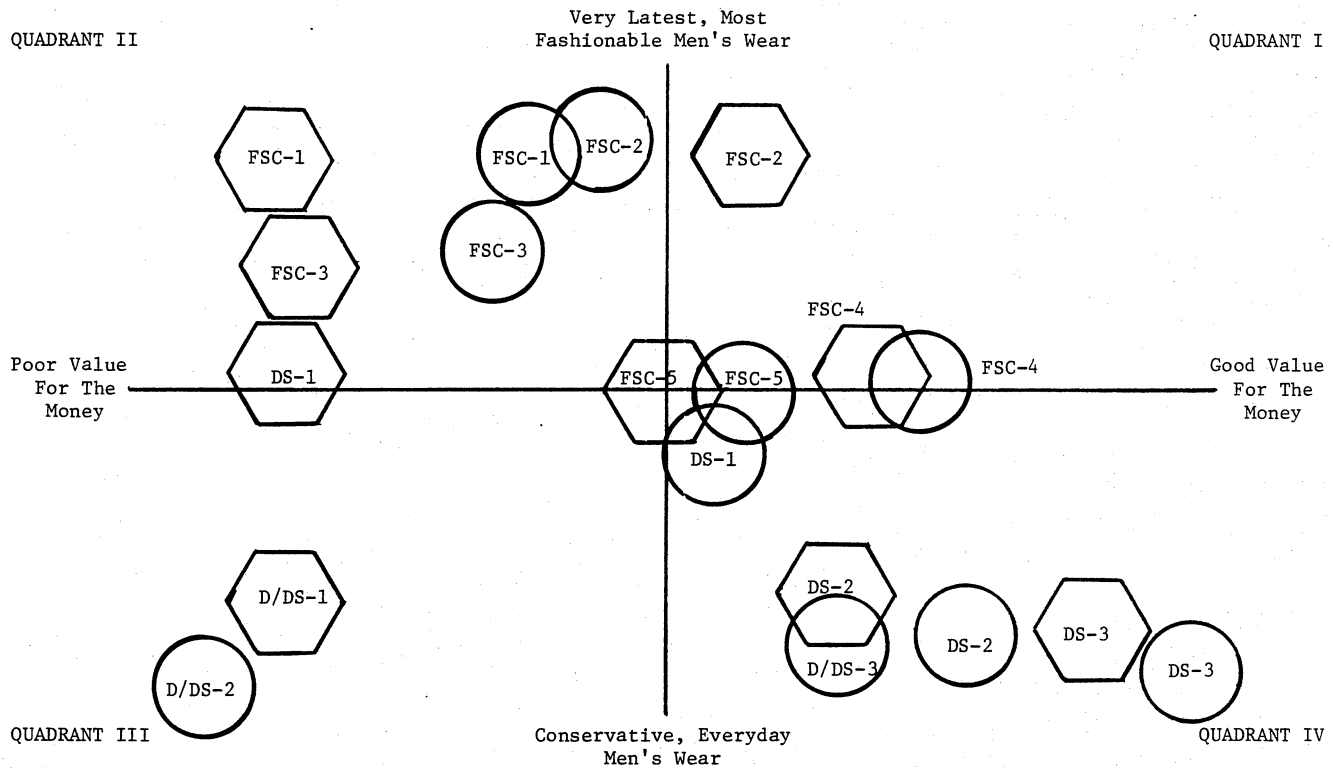
A Two-Dimensional Perceptual Store Image Map

The objective of this section is to present an illustrative example of the output of this local retail market monitoring research procedure. Recall the critical strategic question posed at the outset of this paper: how is the retailer's store image actually perceived by the target customers compared with the firm's competitors in the local market over time?

Toward that goal, a geometric representation--a two-dimensional perceptual store image map--has been prepared. The map presents consumers' perceptions of the major Toronto men's wear retailers on the two retail

³ For a discussion of the evolution of the Toronto Male Fashion Research Program, see King and Ring (1975)

FIGURE 1
 TWO DIMENSIONAL PERCEPTUAL SPACE: STORE POSITIONS ON
 FASHIONABILITY AND VALUE FOR THE MONEY
 1974 - 1975



Legend:

- Hexigon = 1974 Store Position
- Circle = 1975 Store Position
- FSC = Fashion Specialty Chain
- DS = Department Store
- D/DS = Discount Department Store

Note: During 1974-1975, D/DS-1 went out of business in Toronto, and D/DS-2 and D/DS-3 entered the Toronto market.

patronage dimensions of store fashionability and value for the money over time from April-May, 1974 through April-May, 1975.

Store fashionability and value for the money were selected as the dimensions for the perceptual mapping example. These variables were identified as key determinants of retail patronage among consumers in this product category.

Methodologically, consumers were asked to select the store with the "best value for the money in men's fashions" from a pre-selected set of retail stores. Likewise, the consumers were also asked to rank the same retail stores in terms of the store "best for conservative, everyday men's wear", "best for current, up-to-date men's wear" and "best for the very latest, most fashionable men's wear."

The consumer responses on these two dimensions were analyzed for each store to define a specific retailer's store image on those two dimensions for 1974 and again for 1975. The positions of the stores in 1974 are designated by hexagon figures. The positions of the

stores in 1975 are designated by circles. In Figure 1, the 1974 and 1975 maps have been presented as an overlay. The shifts in retailer image and market position on these two dimensions over time are clearly visible.

Strategic Implications of the Perceptual Store Image Map: Retail Market Structure Analysis⁴

At the outset, the two dimensional perceptual map presented in Figure 1 is acknowledged to be a simplistic approach to data analysis. In that simplicity, however, lies its value. The basic concepts involved and the amount and direction of change are graphically presented and are readily understood by the retail manager.

More complex analytical procedures have also been employed on the same data. Ring and King (1977) used multiple discriminant analysis to produce a multi-dimen-

⁴Professor Douglas J. Tigert, University of Toronto and Professor Lawrence J. Ring, University of Virginia were co-researchers in the initial development and analysis of the two-dimensional perceptual store image map.

sional map descriptive of the competitive structure in the Toronto men's wear market. That analysis employs a much more sophisticated computational technique but the results agree very closely with the more simplistic two dimensional map. While the multiple discriminant analysis has more academic appeal, at the pragmatic strategic management level, the more complicated mapping technique may confuse the non-quantitative retailing manager.

Turning to the analysis of Figure 1, for ease of interpretation, Figure 1 can be viewed as being comprised of four distinct quadrants. Each quadrant represents a significantly different retail store image.

Quadrant I in Figure 1, the upper right hand quadrant included those retailers perceived as offering both high fashionability, "the very latest, most fashionable men's wear", and good value for the money. Quadrant II, the upper left hand quadrant, included those retailers perceived as offering high fashionability but poor value for the money. Quadrant III, the lower left hand quadrant, included those retailers perceived as offering low fashionability, conservative, everyday men's wear, and poor value for the money. Quadrant IV, the lower right hand quadrant, included those retailers perceived as offering low fashionability but good value for the money.

Several strategic implications immediately emerge from Figure 1. The overall market structure is well defined in the two dimensional fashionability/value for the money variable space. All of the chains except Discount/Department Store 1 and Fashion Specialty Chain 2 in 1974 and Discount/Department Store 2 in 1975 were positioned in either Quadrant IV, low fashionability and good value for the money, or Quadrant II, high fashionability and poor value for the money.⁵

Note that in both 1974 and 1975, Quadrant I, high fashionability and good value for the money was "under occupied." Only Fashion Specialty Chain 2 was positioned in that Quadrant in 1974 and the Quadrant was empty in 1975. The vacancy in that Quadrant clearly represents a unique market opportunity for store positioning.

The vacancy in Quadrant I, high fashionability and good value for the money, may be the result of various factors. Consumers may see an incompatibility between high fashion and good value for the money. In that case, it would be extremely difficult for a retail chain to develop the appropriate store image and occupy that market position. Alternatively it may be that no existing firm has made an overt strategical attempt to position itself as a retailer of very fashionable merchandise with good value for the money.

In either case, the strategic implication is that there is a potential opportunity in Quadrant I, high fashionability and good value for the money, for either new men's wear retailers or for a shifted positioning of existing retailers.

Looking further at the aggregate market structure, in 1975, the end points of the competitive market space are clearly marked by Fashion Specialty Chains 1 and 2, Department Store 3 and Discount/Department Store 2. Fashion Specialty Chains 1 and 2 clearly lead the high fashionability dimension. Department Store 3 occupies the end point position of the low fashionability and good value for the money store image dimension.

⁵ All of the retailers researched were chain operations with multiple stores. The term store and chain are used interchangeably in the text. The research, however, dealt with the overall perceptual image of the chain.

In 1975, Discount/Department Store 2 dominates the unenviable Quadrant III, poor value for the money and low fashionability. Interestingly, after the 1974 data were reported to the management of Discount/Department Store 1, the firm's management withdrew that operation from the Toronto market, closed or renamed the outlets and launched a new strategical store image development and market positioning program. The critical strategic question facing Discount/Department Store 2 is - is the discounter image of low fashionability with relatively poor value for the money a tenable strategic market position in the Toronto market?

While the end point positions are clearly differentiated, the store images of the retailers in the middle of the map are not well defined. This is particularly true for Fashion Specialty Chain 5 and Department Store 1 in 1975. These firms lack a clarity of offering. Neither of the firms have particularly strong positions on either the fashionability dimension or the value for the money dimension.

By comparison, the positions of Department Stores 2 and 3 have been relatively consistent across 1974 and 1975. Both department stores were well anchored in Quadrant IV, low fashion leaning toward good value for the money though both department store chains modestly improved their value for the money perceptions over the 1974-1975 period.

Adaptive Competitive Strategies by the Individual Retailer Over Time: A Longitudinal Example

The perceptual mapping exercise is also particularly useful in alerting a retailer to changing conditions in the market structure. In response to these identified market position shifts, a retailer can devise an adaptive competitive strategy to respond -- reposition himself.

The strategic positioning of Fashion Specialty Chain 1 over the 1974-1975 period is an illustrative case. In 1974 when the first exploratory, base-line study was reported to the cooperating retailers in the research program, the management of Fashion Specialty Chain 1 was surprised--stunned--at the store image and market position of their firm.

Fashion Specialty Chain 1 was perceived as offering high fashionability but poor value for the money. The firm, in fact, held an end point position of this store image.

In terms of relative competitive posture, Fashion Specialty Chain 1 was an arch rival of Fashion Specialty Chain 2. Fashion Specialty Chain 2, however, had been perceived by the consumer sample to be tied for the top position on high fashionability but with dramatically higher value for the money. Fashion Specialty Chain 2, in fact, was the dominant retailer in Quadrant I, high fashionability and good value for the money.

Over the 1974-1975 period, the management of Fashion Specialty Chain 1 developed and launched an aggressive competitive strategy designed to reposition the firm in relation to its key competitor, Fashion Specialty Chain 2. The strategic maneuver involved an integrated program of assortment modification, consumer advertising, price point realignment, salesclerk training and display merchandise modernization.

The end result of the extensive repositioning effort was reflected in the 1975 local retail market monitoring survey. Fashion Specialty Chain 1 had dramatically improved its value for the money position while maintaining its fashionability image. Likewise, over the ensuing year, Fashion Specialty Chain 2, the key competitor, had slipped from its position of 1974.

In 1975, Fashion Specialty Chain 2 did still hold a modest lead in value for the money. The responsive strategy of Fashion Specialty Chain 1 had been successful in repositioning the firm relative to the total market structure and, in particular, relative to its major competitor, Fashion Specialty Chain 2.

Conclusions

The major strategic problem facing the retailer involves market positioning. The retailer must develop an integrated marketing program to position his firm within his changing competitive environment. Local retail market monitoring is a key element in the retailer's strategic planning.

Local retail market monitoring is an extension of traditional store image research into an on-going, more systematic and standardized research program. Local market monitoring can produce a perceptual store image map of an aggregate retail market structure, locate the individual firm within that map, and identify opportunities unfilled in the market structure. Likewise, local retail market monitoring can track changes in the firm's position over time as a result of competition or as a result of purposive strategy changes by the retailer.

References

Charles W. King and Lawrence J. Ring, "The 1975 Toronto Male Fashion Research Project: A Descriptive Overview," Institute for Research in the Behavioral, Economic, and Management Sciences, Purdue University, Paper No. 526, 1975.

Jay D. Lindquist, "Meaning of Image," Journal of Retailing, 50(Winter 1974-1975), 29-38.

Eleanor G. May, "Practical Applications of Recent Retail Image Research," Journal of Retailing, 50(Winter 1974-1975), 15-20.

Pierre Martineau, "The Personality of the Retail Store," Harvard Business Review, 36(January-February 1958), 47-55.

Lawrence J. Ring, "The Male Fashion Consumer: An Analysis of Fashion Involvement and Retail Patronage Determinants," Unpublished Doctoral Dissertation, Purdue University, West Lafayette, Indiana, 1977.

Lawrence J. Ring and Charles W. King, "A Multiple Discriminant Analysis Approach to the Development of Retail Store Positioning," Advances in Consumer Research, Vol. V, (1977, Proceedings).

Robert G. Wyckham, William Lazer, and W. J. E. Crissy, "Images and Marketing: A Selected and Annotated Bibliography," (Chicago: American Marketing Association, 1971).

CROSS-CULTURAL RESEARCH METHODOLOGY
AS A CASE OF CONSTRUCT VALIDITY

W. Fred van Raaij, Tilburg University

Abstract

Cross-cultural consumer research witnesses an increasing interest of researchers, both for the managerial strategy of multinational companies and in order to establish the universality or specificity of theoretical construct and findings.

In this paper the validity of the hypothetical constructs is discussed in order to create functionally equivalent constructs across cultures and equivalence of samples from the cultures.

A theoretical structure in which a hypothetical construct is embedded serves as a manner for attaining the functional equivalence of instruments and constructs in order to interpret the findings and to implement managerial strategy.

Introduction

In 1889 Sir Francis Galton presided over a meeting of the Royal Anthropological Institute of Great Britain in which a paper was read using cross-cultural data. At the meeting Galton criticized this paper, raising the problem of independence of cases. Elder (1976) describes this as follows:

"If two societies exhibit the same juxtaposition of institutions or traits, how does one know whether these two societies represent two independent cases of that juxtaposition? What if society A "borrowed" the juxtaposition from society B, in which case the juxtaposition is historical rather than functional? What if A and B are merely variants of some common overarching society C? In either case, what appear to be two cases may actually be two illustrations of the same case, in which event the cross-cultural comparison has not produced any further evidence for the generalization of the juxtaposition than did the single-culture observation."

This illustrates the theoretical value of cross-cultural research: the generalization of findings and constructs across cultures, or better, following Popper (1959), the attempts to invalidate a theory or "behavioral law" as culture-specific.

From a practitioner's viewpoint, there exists the need for information on consumer behavior in other cultures and nations in order to develop international marketing operations. It's not the validation or invalidation of theories that is important here, but rather the answer on the question of applying a global or culture-specific marketing strategy.

Triandis, Malpass, and Davidson (1973) mention the following reasons for conducting cross-cultural research:

"To check the generality of behavioral laws; to increase the range of our observations on variables of interest; to determine the variations found in subjective culture variables in different settings; to take advantage of natural experiments involving combinations of variables (conditions) that cannot be obtained in the laboratory; to study the manifestation of psychological variables in different

cultural contexts; and to study cultures for their own sake."

Such research can serendipitously lead to a better insight into the communalities and specificities of consumer behavior in different cultural settings and the universality of behavioral "laws" and concepts across cultures (Poortinga, 1975).

Some introductory remarks make clear the ground for the remainder of this paper:

First, cross-cultural research as comparative research is not unique. All behavioral research is comparative; that is, it involves a comparison of experimental and a control group, e.g., samples of working and non-working wives, husband and wife roles, middle and lower social class consumers, users and non-users of a service, etc. Cross-cultural research is unique in that it considers different cultural settings or "geographical environments" (Koffka, 1935) that shape consumer behavior and are shaped by consumer attitudes and behaviors. Therefore, it means that we cannot control for the "disturbing" factors that may create unwanted differences between the samples under study, and it means that we cannot use the same instrument (questionnaire) for both samples in different cultures.

Secondly, consumer research is largely "made in the USA." Concepts and instruments have been developed in the United States and not in the "other" culture of the cross-cultural study. This may introduce "ethnocentrism" in the type of questions we address, the concepts we employ, and the explanations we give of the results. For instance, the study of "consumer satisfaction" is relevant in a western mass-consumption society but not, or in a different way, in a developing country. Another example is a book by Nieuwenhuijze (1963) reporting the results of an analysis of Islam, that was banned in Indonesia on the grounds that a non-Muslim could not properly examine either the history or the tenets of Islam. A last example is a questionnaire statement: "A good citizen is responsible for shoveling the sidewalk in front of his home" for measuring "social responsibility." This statement assumes private ownership of houses, one-family housing, and a climate with snow in winter, and is clearly not applicable in an African country. In that case, we have to measure "social responsibility" with another statement or set of statements.

Thirdly, the terms "cross-cultural" and "cross-national" are used in the literature. I prefer "cross-cultural" because this term reflects more possible differences in consumer behavior than "cross-national." In cross-cultural research we compare consumer behavior in different cultures, including subcultures, e.g. differences between French- and English-speaking Canadians; black and white subcultures; and minority ethnic groups with the majority ethnic group. Many of the problems and caveats that apply for cross-cultural research are also valid for market segmentation studies using the segmentation variables of race, ethnic background, and social class.

Cross-cultural studies on consumer behavior include: shopping behavior (Green and Langeard, 1975; Douglas, 1976), innovator characteristics (Green and Langeard, 1975), female role perception (Douglas, 1976), media

usage (Douglas, 1977), family buying decisions (Hempel, 1974), life-style research (Plummer, 1977), and information seekers (Thorelli, Becker, and Engledow, 1975; Anderson and Engledow, 1977; Becker, 1976). These studies compare a US sample of consumers with samples in France, England, Canada, Mexico, and Germany, all nations of the western world. Sheth and Sethi (1973) developed a theory of cross-cultural buyer behavior for the diffusion of innovations, and Dubois (1972) developed a framework for the study of the cultural factors that affect the rate of adoption of an innovation. Green and White (1976) give methodological considerations in cross-cultural consumer research. De Vos and Hippler (1969) and Triandis, Malpass, and Davidson (1973) review cross-cultural research in psychology.

History

Cross-cultural research is not new. A long cross-cultural tradition exists in anthropology. Benedict (1934) suggests that cultures - especially of small, isolated groups - are often integrated wholes wherein the parts coalesce around certain basic values. To extract parts from the wholes, compare them (out of context) with parts extracted from other cultures, and then conclude that one has found cross-cultural similarities or differences does gross injustice to social reality. Osgood (1967), using the semantic differential technique, demonstrated that persons speaking different languages have different subjective cultures. Arguments still continue over the degree to which language determines thought processes and thus subjective culture: The Whorfian Hypothesis (Whorf, 1956). The argument in anthropology is that cross-cultural data can be used primarily for the identification of cultural uniqueness or the establishment of cross-cultural contrasts, describing the contrasting cultures in their own context. However, in consumer research, our interest is not primarily the Eskimo of Northern Canada or the Temne of Sierra Leone (Berry, 1967). We study more-or-less similar cultures in Northern America or Western Europe. We may expect more culturally unique behavior patterns in the study of Japanese consumers or consumers in developing countries.

Borrowing the psycholinguistic terminology, a distinction will be made between an emic (culturally specific) and an etic (culturally universal) approach. In an emic approach (from phonemics) the behavior is described in terms and concepts of that specific culture and with internal criteria; hence cross-cultural comparison is difficult or impossible. In an etic approach (from phonetics) the behavior is described in universal categories and with external criteria. Cross-cultural comparison is feasible in the etic approach. Davidson (1977) criticizes the "emic-etic dilemma" and advocates the use of etic concepts (with which comparisons among cultures can be made) and emic ways of measurement (in the context of the culture under study). We will see that this is the central issue in cross-cultural methodology, also referred to as the functional equivalence of measuring instruments.

Another historical development is the evolutionary or revolutionary theory of socio-economic change. Cross-cultural contrasts are used to identify, in a quasi-experimental way, why socio-economic changes in one culture differ from those in another culture. Weber (1958) discusses why industrial capitalism developed in western Europe but not in China or India (Protestant ethic). Certain cultural conditions were different or absent in China or India, and industrial capitalism did not develop there. Hsu (1963) explained why clan, caste, and club emerged respectively in China, India, and the

United States. Implicit in Hsu's analysis is the counterfactual premise that if family structure had been different in the USA, clans or castes rather than clubs might have emerged.

Marx (1972) argued for a contextual view of social phenomena; the basic mode of production in a culture or nation determines social and cultural phenomena. Engels (1949) identified a sequence of politico-economic stages based on the means of production and the class relationship to the production means. Knowing these stages (feudal, bourgeois-capitalist, socialist) one can develop generalizations for societies in the same stage. Parsons (1966) distinguishes primitive, advanced primitive, archaic, advanced intermediate, and modern societies, identifying certain prerequisites as necessary but not sufficient conditions for a society's transition from one stage to another. Meaningful cross-cultural comparisons can only be made between societies at the same level of development (differentiation). Consumer behavior differs depending on whether it occurs in a capitalist nation, a dependent nation (colony) or a socialist nation.

Sampling

Cross-cultural research involves the comparison of samples from different populations (cultures). What type of sample we select depends on our research objective.

1. For a comparison of income and age distributions in different cultures one needs a random sample to employ statistical tests for the significance of differences.
2. For descriptive studies on attitude-value structures, attitude-behavior relationships, life-style, or opinion research sample representativeness is essential. This means that the sample has to be representative for the culture from which it is drawn. In this case, we study the relationships between variables in different cultures and we are not primarily interested in the absolute scores (distributions) of the variables as such.
3. In some cross-cultural studies we need functional equivalence of samples, e.g., comparing salesmen in an American and a Japanese department store. Here we try to find "similar" organizations and samples of the personnel of the organizations. Generalization of the results to the general culture, however, becomes hazardous.
4. For causal studies across cultures, we may employ "matched" samples with matching variables such as age, income, education, etc. One increases the power of the statistical test for the significance of differences using paired observations. But the danger exists that one eliminates real cross-cultural differences through the matching procedure. One has to have a theory to distinguish between the matching variables and the independent variables in a research design.

Except for the random samples, sampling procedures involve an a priori distinction between variables "that are controlled for" in sampling and "real" independent variables that are important from a theoretical viewpoint. To control for sampling inadequacies and to partial out covarying factors that could not be controlled for in the sampling procedure, the technique of analysis of covariance (ANCOVA) may be used. But one has to be aware of the dangers of these control procedures. Controlling for obvious variables such as age, income, and education in the sampling procedure may

create samples unrepresentative of the population. A sample from a culture in India with the same age, income, and education distribution as a US sample may be largely unrepresentative of the Indian culture.

Pretest data will often help to pinpoint the best sample. Pearlin and Kohn (1966) were interested in sampling people from equivalent class positions in the USA and Italy. Their preliminary data showed that income and education were not suitable for manifesting similar class positions, but that occupational prestige was. They approached the sampling problem by minimizing the extremeness of the groups chosen to be interviewed, finally selecting members of the middle and working classes. They concluded that "some intra-class variation is obscured by using only those two broad social class categories but what is lost in precision is gained in increased cross-national comparability" (p. 468).

Brislin, Lonner, and Thorndike (1973) advocate the plausible rival hypotheses approach. The differences or similarities that are found may be attributed to different sampling methods or to different qualities of the samples (age, socioeconomic status) rather than to "real" cross-cultural differences or similarities. A rival hypothesis may explain the obtained differences/similarities as well as the hypothesis one is interested in. In the research design one has to eliminate all plausible rival hypotheses that may explain the results from sampling inadequacies. It may be clear that "identical" samples in cross-cultural research are actually impossible; that the costs of random samples are often too high to justify their additional benefits. Essentially, equivalence of samples, based on the control of variables in a design that eliminates plausible rival hypotheses, must be the objective of sampling in cross-cultural research.

Rival hypotheses cannot always be ruled out in a design with only one sample from each culture. If we study the differences and similarities of French and American wives, it may be illuminating to include control samples of French and American husbands. How do the wives differ from the husbands (not necessarily their husbands!) in the two cultures, and an interaction (in the statistical sense) may exist between sex and culture. A good example is the 2 x 2 design by Douglas (1976), where she compares working and non-working wives in France and in the USA, using multivariate ANCOVA. Cross-cultural differences in Douglas' study prove to explain more variance than the differences between working and non-working wives.

Thorelli, Becker, and Engledow (1975), on the other hand, conclude that information-seeking consumers (higher education and higher income) are more similar across cultures (the USA and Germany) than the control group of non-information-seekers. Katona, Strumpel, and Zahn (1971) find more similarity across nations for white-collar than for blue-collar respondents.

Formal and Functional Equivalence

Formal equivalence of a questionnaire or a questionnaire item means that an identical questionnaire and respondent instruction are employed for all respondents involved in the study. Formal equivalence is the ideal of the survey technician but is impossible in cross-cultural research in nearly all cases. Translation problems and procedural problems such as whether to use a mail questionnaire or a personal interview, may arise. Formal equivalence can be attained to some extent through back-translation (translating the questionnaire and

instructions into the other language and then back into the original language by independent translators, and then comparing the two versions), the use of bilingual respondents, expert judges, and writing translatable English. See Brislin, Lonner, and Thorndike (1973, pp. 32-81). Formal equivalence is treated as an instrument reliability problem: How reliably does the instrument measure the variables in different cultural settings? It will be clear that we can easily attain formal equivalence in measuring behavioral variables (e.g., possession of durables; number of hours of TV watching; frequency of shopping trips). In these cases, we are interested in measuring and comparing behavioral variables as such. But as soon as these behavioral variables become an indicator or operationalization of a hypothetical construct, it becomes another case. "Frequency of church attendance" may be selected as an indicator of the hypothetical construct "religiosity." In this case, church attendance may be a good indicator of religiosity in one culture but not in another.

Functional or conceptual equivalence of instruments for measuring a certain construct is a validity problem. The survey methodologist strives for functional equivalence of instruments, i.e., the instrument must have a functionally equivalent meaning for the respondents in both cultures. Although questionnaire items may be different for different cultures the intention is that they measure the same hypothetical construct. "Religiosity" can be operationalized by church attendance in one culture while meditation may be an indicator of religiosity in another culture. In such cases, functional equivalence of instruments can only be attained through a theoretical a priori framework and pilot studies how a hypothetical construct becomes manifested in actual behavior or attitudes in the culture under study.

Straus (1969) distinguishes formal and functional equivalence of stimulus material (questionnaire, test, task) and of mode of quantifying. Instead of the terms "formally and functionally equivalent" he uses the terms "phenomenally identical" and "conceptually equivalent." Straus (1969) obtains the four possibilities of Figure 1, using our terminology.

Figure 1

A Taxonomy of Measurement Equivalence

Mode of Quantifying	Stimulus Material	
	formal equivalence	functional equivalence
formal equivalence	culturally universal (etic)	culturally modified
functional equivalence	culturally ipsatized	culturally specific (emic)

The mode of quantifying involves the development of norms for coding and scoring. A psychological test can only be used for a new population provided new norms are developed based on standardizing the test among samples of that new population (culture). The four quadrants have the following interpretation:

1. Culturally universal measurement refers to research methods that use identical stimulus materials in each culture, except for translation, and that record and

quantify the responses in identical ways. This means the formal equivalence of instrument and quantification of responses. This is the "etic" approach to cross-cultural research. It is likely that there are only a limited number of instances. The meaning and intention of this observable behavior, however, may be different across cultures.

2. Culturally ipsatized measurement refers to instances in which the identical instrument is used but the recording and/or interpretation of the responses are judged relative to others in that culture (internal criteria) rather than relative to external standards of scoring or interpretation. An attitude scale measuring "ecological responsibility," for example, may be identical across cultures but the norms for high or low scores may differ across cultures.

3. Culturally modified measurement refers to measurements in which the indicators for the hypothetical construct are altered in order to make them culturally appropriate but the original scoring is maintained. Przeworski and Teune (1966/67) provide an example in their cross-cultural study in the U.S.A. and Poland. The hypothetical construct "political activity" is measured in the U.S. with the indicators "contribute money to parties or candidates" and "place sticker on car," but in Poland they used the equivalent indicators "fight for the execution of economic plans" and "attempt to influence economic decisions." Przeworski and Teune (1966/67) also had a set of five identical indicators for both cultures.

4. Culturally specific measurement in the fourth quadrant represents the minimum formal equivalence in order to achieve the maximal functional equivalence. This is the "emic" approach to cross-cultural research, although it does not state the absolute uniqueness of cultures (Benedict, 1934). To measure the extent to which a sample of youth participate in an adolescent subculture, items of dress reflecting adolescent fads of the particular subculture might be used as a part of the index. In one subculture such indicators may be tight pants and in another subculture bell-bottom pants.

Straus (1969) provides four techniques for establishing functional equivalence of measures:

1. Use of expert judges to evaluate the appropriateness of the items, questions, and instructions. This procedure may help provide for the content validity of the instruments.
2. Revalidation and restandardization of the instrument in the culture in which it will be employed (ipsatizing).
3. Semantic differential procedure to obtain "psychological translation" of indicators. Blood and Takeshita (1964), as cited by Straus (1969), found that the word *deito*, a translation of the English "date", was not widely enough known in Japan to use it in the questionnaire. They therefore substituted the work "otsus ia" with connotations similar to "date."
4. Construct validation. Construct validity is determined by establishing that certain hypothetical constructs account to some degree for performance on a psychological test, or that the test accounts to some degree for performance on variables for which the test is taken as explanatory (APA standards). Zaltman, Pinson, and Angelmar (1973) define construct validity as "the extent to which an operationalization measures the concept (hypothetical construct) which is purports to measure." For instance, Blood and Takeshita (1964) measured the relative power of husband and wife in the Japanese family. The item "Who decides about what

type of car to get?" had to be eliminated and had to be substituted by some major purchase. Formal equivalence is given up for the sake of functional equivalence. The construct validation of such an instrument would consist of determining if the scores are related to other variables which theory or prior knowledge indicated should be related (Straus, 1969):

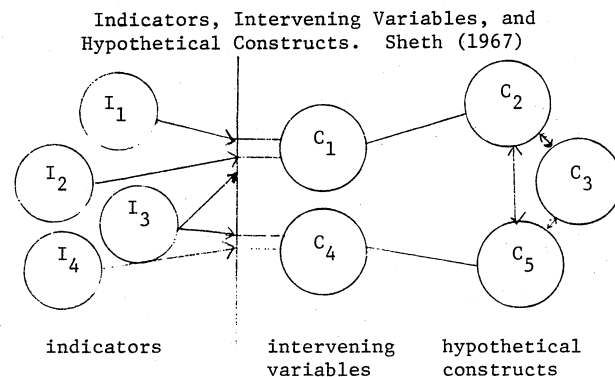
Another approach to functional equivalence of instruments is provided by Przeworski and Teune (1966/67) distinguishing between identical and equivalent indicators. An identical indicator is universal (etic) across cultures and an equivalent indicator is specific (emic) for a culture. We have seen the example of equivalent indicators in their study on political activity in the U.S.A. and Poland. By combining identical and equivalent indicators reliable and valid comparisons can be made. By this procedure a longer and more culturally relevant, and perhaps more valid, set of items is acquired. A set of items for a culture consists of universal and specific items. Inter-item correlation reveals that the specific items add to the universal items.

Green and White (1976) make a distinction between the functional equivalence of instruments, of phenomena ("shopping"), and of constructs ("Cognitive consistency"). However, an instrument is functionally equivalent across cultures if it measures the "same" construct.

Theoretical Structure

Construct validity is not only the extent to which an operationalization (set of indicators) measures the hypothetical construct which it purports to measure. It is also the relationship of the hypothetical construct with other hypothetical constructs and intervening variables in a theory. A theory is a set of propositions relating hypothetical constructs, formulated in such a way that testable hypotheses can be derived (de Groot, 1969). Maehr (1974) gives an example of the relationships of the hypothetical construct "achievement motivation" and other hypothetical constructs. Cross-cultural studies of "achievement motivation" suggest that the psychological dynamics underlying achievement behavior (e.g. entrepreneurial action) might not be the same across cultures. In other words, variables which

Figure 2

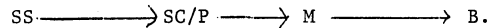


motivate individuals to be economically productive in one culture may differ from those motivating productive

work in another culture. These variables may be the need for affiliation, concern for the reaction of others, and strong obligations and identifications to one's group. In Japan the basic cultural value is

loyalty to one's group rather than the advancement of one's personal status. The hypothetical construct "achievement motivation" is conceptually different across cultures and even irrelevant in some cultures. We have to specify when and under what conditions achievement motivation will occur. Maehr (1974) proposes three theoretical structures in the study and prediction of achievement motivation.

1. In the first structure, social structure (SS) or, more specifically, the social learning experience provided by the milieu in which the person develops, influences the subjective culture (SC) or the assumed predispositions (attitudes, values, norms) and the personality traits (P) to respond in a given manner. SC determines achievement motivation (M) and achievement behavior (B). In a diagrammatical expression this may be indicated as:



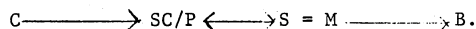
This is the traditional approach by McClelland (1961) where the personal history of the individual determines his achievement motivation. Only through changes in SC/P achievement motivation may be changed.

2. A second structure may be diagrammed as follows:



where a situation or context (S) has demonstrable effects on motivation (M). Subjective culture (SC) and personality (P) are placed in parentheses indicating the relative unimportance of the attitudes, norms, values and personality variables. Here situational variables play a major role in determining achievement motivation. Maehr (1974) identified four situational dimensions that appear to be critical in determining achievement motivation and behavior. An individual achieves as a member of a social group and is influenced by the norms of his group. Another dimension is the locus of control. How much (internal) control has the individual over the outcomes (success or failure) of his behavior? Feedback about performance has to be communicated to the achiever through interpersonal contacts. A task dimension involves task difficulty and intrinsic task interest. The situational model implies that the above situational variables greatly affect achievement motivation and behavior. It is better changing the situation rather than the personality of the individual to enhance achievement motivation.

3. The third theoretical structure represents the most sophisticated analysis of achievement motivation:



This structure identifies both subjective culture/personality (SC/P) and situation (S) as simultaneously critical variables. This SC/P x S model goes back to Lewin; one of the prominent and highly developed of such models is the motivation model proposed by Atkinson (1964, Chapter 10). See also Van Raaij and Wandwossen (1978). The question becomes to find what situations match what motives in order to elicit a specific behavior such as achievement motivation. This theoretical structure requires considerable theoretical and methodological sophistication.

4. Not proposed by Maehr (1974) but as a counterconceptualization, Inkeles (1971) states that personality structure does not necessarily precede modernization and economic development. The influences may go in the opposite direction. That is, institutional changes may change the outlook of individuals exposed to them. According to Inkeles' theory modern social and cultural structures and systems may cause individuals' attitudes,

values, behaviors, and need dispositions to change in such a way as to make effective adaptation to these institutions by becoming increasingly "modern." The greater the degree of exposure and the longer the exposure time to the various modern institutions, the more the individual will shift toward the modern

direction on the psychological continuum. Inkeles' theory offers an alternative way for emphasizing the role of the individual in social institutions. The diagram becomes as follows:



The physical environment (PE) and/or the social structure (SS) expose an individual to situations (S) of modernization where achievement motivation (M) and behavior (B) are the best adaptive responses. M and B change the subjective culture (SC) and personality (P) of the individual afterwards making it consonant with behavior.

These four theoretical structures have consequences for the measurement and change of achievement motivation. In the first structure, projective personality tests have been used, and we may change M and B through changes in SC/P (personality change and attitude change). In the second theoretical structure, M and B are determined by situational dimensions. Group norms, locus of control, feedback, and task dimensions are indicators of M and B, and we may influence M and B by changing the situational dimensions. The interaction of SC/P and S determines M and B in the third structure. In the fourth structure, achievement motivation is the result of environmental factors (PE and SS), and is in fact adaptive behavior to the changing (modernized) environment. Personality and attitude (SC/P) measures are not a good predictor of M and B in a changing environment but the situation (S) itself may become an indicator of achievement motivation (M) and behavior (B). Maehr (1974) advocates a methodology of identifying settings in which achievement motivation occurs or does not occur in various cultures and then proceeds to characterize the nature of these settings. Much of the research that attempts to understand motivational patterns of (sub)cultural groups involves placing individuals in a "middle-class biased" or "western-biased" performance setting and then observing behavior. But instead, we have to devise a theoretical structure of achievement motivation for the culture under investigation and from this theoretical structure functionally equivalents may be derived for measuring the hypothetical construct of achievement motivation.

Cross-cultural Consumer Research

Triandis, Malpass, and Davidson (1973) distinguish classes of distal variables (PE, SS, O) which act as independent variables in studies in which classes of proximal variables (A, SC, P, B) are the dependent variables. These classes of variables are for consumer research:

- PE: physical environment: type of economy, distribution structure, climate, population density, GNP, transportation system, consumption level, health and education system. Sethi (1971) uses PE variables for clustering nations into eight groups.
- SS: social structure: social class, family organization, social and sex roles, social norms, language, socialization and education patterns. Sheth and Sethi (1973) call the PE and SS variables exogenous and treat them as givens or constraints of their model. These variables are not explained in terms of their structure or any changes over time.
- O: other persons: influence attempts, opinion leadership, social interaction, word of mouth, imitation, conformism/deviancy, cooperation/competition.

- A: persons' abilities: intelligence, cognitive style, cognitive development, memory, information-processing capabilities.
- SC: person's subjective culture: attitudes, values, personal norms, aesthetic judgments, opinions, life style, beliefs and stereotypes, meaning, satisfaction.
- PD: personality dispositions: self-concept, anxiety, aggression, alienation, achievement and affiliation motives, adjustment, cosmopolitanism, information-sensitivity.
- B: behavior: shopping, communication, consumption, and disposition behavior.

In cross-cultural consumer research, as indicated earlier, we cannot simply compare concepts from a class across cultures. We have to study the functional equivalence of these concepts (hypothetical constructs) across cultures through both their operationalization (indicators) and their relationship to other concepts in a theoretical structure. Maehr (1974) and Inkeles (1971) already provided four possible theoretical structures for the study of achievement motivation and economic behavior.

I will try to give some examples in consumer research of possible theoretical structures.

- The cross-cultural study of consumer satisfaction has to take into account the PE variables: type of economy, distribution structure, consumption level, transportation system, and income distribution, and the SS variables: social class, family organization, social and sex roles. If the cross-cultural differences of the PE and SS variables are great it becomes meaningless to compare consumer satisfaction scores. Consumer satisfaction is a function of the PE, SS variables and probably also of O and A variables.

PE → SS → A/O → SC

Consumer satisfaction as such may become functionally equivalent across cultures that differ on PE/SS variables but we have to employ a different (emic) set of indicators for each culture. Comparing the consumer satisfaction of a mass-consumption society with a developing country becomes a "tour de force" in finding sets of equivalent indicators. In a case of extreme differences, we better make cross-cultural comparisons of PE/SS variables only. If only slight differences on PE/SS variables of two cultures exist we may find sets of equivalent indicators in both cultures for the hypothetical construct "consumer satisfaction." The procedures may be largely emic in that they reflect the PE/SS structure in a culture but may become "psuedo-etic" in that a comparison is made between similar evaluative consumer reactions to the consumption system.

- Douglas (1976), Green and Langeard (1975) compare among other variables grocery shopping behavior in France and the U.S.A. Shopping behavior (B) can be thought of as a function of the PE variables: transportation system, private ownership of cars, distribution (retailing) structure, consumption level, and the SS variables: family organization, social and sex roles, working or nonworking housewives. Intervening variables from the O, SC, and PD classes may influence shopping behavior: social interaction, word of mouth, life style, attitudes, self-concept, and opinion leadership. Two theoretical structures can be hypothesized:

I: PE → SS → (O/SC/PD) → B,

II: PE → SS → B → (O/SC/PD).

PE/SS variables influence the O/SC/PD group, and O/SC/PD determines behavior, or PE/SS determines behavior and the O/SC/PD variables change into a direction to become consistent with behavior. Green and Langeard (1975) conclude that "most of the differences that we found (between samples of French and U.S. women) could be attributed to social and environmental that characterize the two nations. These finding emphasize the need for further theoretical and empirical cross-national work to achieve a better understanding of the relationships between buyer behavior and environmental forces." Douglas (1976) concludes in a similar fashion: "Retail environmental factors seem to play a key role in shaping and conditioning behavior patterns between countries may in many respects merely reflect current market conditions such as the availability of different products and services or the number of supermarkets or small, traditional retailers." She recommends further investigation of the role of retail environmental factors in influencing customer response patterns. Shopping behavior does not reflect "real" cross-cultural differences but reflects only the different distribution structures in France and in the U.S.A. Green and White (1976) add that shopping is not functionally equivalent in France and in the U.S.A. "In France shopping is reputed to be an integral part of the housewife's social life, whereas in the United States shopping tends to be considered a chore." Even a third theoretical structure is not ruled out, namely that the distribution structure reflects the social interaction patterns, the subjective culture, and personality dispositions of the customers:

III: (O/SC/PD) → (SS) → PE → B.

Does the "adaptive consumer" adapt his behavior to changes in the distribution structure (model II), or do we have to change the O/SC/PD cluster too before new shopping patterns develop? (model I). Or, do "adaptive retailers" change the distribution structure in line with changes in the O/SC/PD cluster? Most retailers want us to believe that model III is true because it reflects the marketing concept.

- Hempel (1974) found "a surprisingly high degree of cross-cultural similarity in household decision-making processes. In most instances, the differences between the roles perceived by husbands and wives within the same cultural setting were greater than the differences between cultures for either sex." Hempel compared samples in Hartford, Connecticut, and Preston-Lancaster, England, with nearly identical data collection methods, except for some minor adaptations to English word usage. Husband-wife interaction is a function of the SS variables: social class, family organization, social and sex roles, socialization and education patterns. A simple theoretical structure can be postulated:

(PE) → SS → O.

The United States and the United Kingdom are similar cultures in many respects. A valid cross-cultural comparison can be made and functionally equivalent instruments (no translation problem) are easily developed.

- Beginning with the pioneering study of Horton (1943), social scientists have attempted to find out why people drink alcohol. A correlation has been found between a PE variable, type of economy (hunting, fishing vs. agricultural), and a B variable, alcohol consumption. In general, hunting tribes drink more heavily than tribes that depend on agriculture:

PE → SS → (PD) → B.

But a number of variables may intervene the PE → B relationship. The degree of societal organization has received the strongest support as an intervening variable. In more highly organized cultures there are more social controls that inhibit drinking. The reflection of social organization on the individual disposition (PD) has been called alienation. The greater the alienation the more alcohol is consumed.

A recommendation for research designs is that if two hypothetical constructs are not comparable across cultures, i.e. are not functionally equivalent, the researcher needs to go to the next-higher order construct in the theoretical structure. If consumer satisfaction is an incomparable construct, compare the distribution system and/or the social system across cultures.

Construct Validity

Construct validity is the extent to which an operationalization (indicator) measures the construct which it purports to measure (Zaltman, Pinson, and Angelmar, 1973). Or alternatively, in order to show that an indicator applies to a construct, it is necessary to derive hypotheses from the theory related to the construct, that can be tested in reality (indicators) (Cronbach, 1964, p. 105). Thus, we have to include in the operationalization of a construct the theoretical structure in which the construct is embedded.

The validity of constructs across cultures can be attained through four essential procedures:

1. The operationalizations (measuring instruments, indicators) should measure the construct which they purport to measure. If identical indicators (etic measures) are not available, a set of identical and equivalent indicators (Przeworski and Teune, 1966/67), or equivalent indicators (emic measures) should measure the hypothetical construct.

2. The validity of a construct can be established in some cases through systematic variation with a criterion. Criterion-related validity is concerned with prediction (predictive validity) or with simultaneous performance of some other variable (concurrent validity). For instance, we measure the concept of "innovation-proneness" in several cultures using functionally equivalent indicators. The degree of adoption of an innovation may be the criterion measure. This means that we can validate our measuring instrument for the hypothetical construct only afterwards.

3. A hypothetical construct has relations to other constructs in a theoretical structure. Examples are provided in an earlier paragraph. Causal explanations for the obtained cross-cultural differences and similarities are proposed in a theoretical structure.

4. A last essential procedure of construct validity is the similarity of components of the construct across cultures. Frijda and Jahoda (1966) call this "dimensional identity." "Intelligence", for instance, consists of several components (verbal, numerical, etc.) obtained after factor-analyzing the concept. "Intelligence" is functionally equivalent across cultures if the same components are found in different cultures. "Perceived risk" may differ in its components (social, physical, financial risk, etc.) across cultures but that means that the construct validity of the hypothetical construct "perceived risk" is low. Especially in cross-cultural assessment (test) psychology, the componential similarity of a construct is stressed. However, we may study the components of "intelligence" in another culture with an emic approach, i.e. investi-

gating the unique componential structure of a construct in that culture.

Ethnocentrism

One of the biases in cross-cultural research has been called "ethnocentrism", i.e. bias arising from blindness for the unique and different characteristics of another culture. Consumer research is largely "made in the U.S.A." with all the risks that western-American- or middle-class biases pervade this type of research in the research questions we address, the concepts and theories we use, and the interpretations we give. We should encourage researchers in other cultures to study their own reality rather than to replicate American studies. Faucheux (1976) recommends a separate development of social psychology on the European continent. We may recommend the same for consumer research in order to become aware of the ethnocentrism and the blindness to other cultural values. A rationalistic tradition in Europe encourages theoretical speculation, sometimes at the cost of a proper care for empirical validation. This tradition produced scientists such as Freud, Lorentz, Piaget, Levi-Strauss, and Chomski, and does not have the illusion that a theory will spring from a gathering of "facts" and measurements collected with a preconceived idea. Some suggestions for avoiding ethnocentric biases are:

1. Double investigation: The researchers in a cross-cultural study perform their research or experiment in both cultures and compare their results.
2. Decentering of indicators: Questionnaire items are developed in both cultures are then an identical set and two equivalent sets are formed. See Brislin, Lonner, and Thorndike (1973, pp. 37-39).
3. Select and/or train the staff members in the other culture to introduce their own questions, concepts, and interpretations.

Galton's Problem and Evolutionary Theory

How do we get subclasses of cultures in which a valid cross-cultural comparison is possible? And if we want to generalize a finding, how do we get two or more cultures that are not simply two or more illustrations of the same case? This is Galton's problem. Evolutionary theory stated that only cultures in the same stage of differentiation can be meaningfully compared. In any case, we have to classify cultures or nations into diverse subclasses or types. Sethi's (1971) cluster-analytical approach may serve this purpose on a global scale. More refined classification is needed within the western cultures. To a certain extent, we may assume that concepts are functionally equivalent within such a subclass. Galton's problem, however, is to find mutually "independent" cultures (in different subclasses) to make universal generalization of a phenomenon possible, if one's hypothesis proves to be true for these independent cultures. We cite Elder (1976, p. 217):

"Galton's problem bears an interesting relationship to the problem raised by critics of universal cross-cultural generalizations, to the effect that one can meaningfully compare only cultures rather similar to each other in terms of size, structural complexity, etc. In both cases, the methodological goal is to identify cultural subsets to which specific cultures can be assigned. In the case of Galton's problem, the purpose of assigning cultures to subsets is to try to guarantee subsequently that cultures being compared

are drawn from different subsets, thereby permitting universal cross-cultural generalizations. In the case of the critics of such generalizations, the purpose of assigning cultures to subsets is to try to guarantee subsequently that cultures being compared are drawn from the same subset, thereby permitting limited but meaningful cross-cultural generalization."

Conclusion

From a theoretical viewpoint mere comparison of consumer responses to questionnaire items in different cultures does not guarantee meaningful and useful information for theoretical purposes and for managerial action. Functional equivalence of constructs and instruments has to be established in order to guarantee "real" cross-cultural comparison. Plausible rival hypotheses that also explain the obtained differences/similarities have to be ruled out by using equivalent samples, and functionally equivalent measures and quantification modes.

Cross-cultural research as comparative research creates some additional problems because of the non-equivalence of instruments that have to be employed. Only through (1) equivalence of indicators, (2) criterion-related validity of constructs, (3) a theoretical structure, and (4) componential similarity of constructs can the validity of hypothetical constructs be determined.

References

- Ronald Anderson and Jack Engledow, "A Factor Analytic Comparison of U.S. and German Information Seekers," Journal of Consumer Research, 3(March, 1977), 185-196.
- John W. Atkinson, An Introduction to Motivation, New York: Wiley, 1964.
- Helmut Becker, "Is there a Cosmopolitan Information Seeker?" Journal of International Business Studies, 7(Spring, 1976), 77-89.
- Ruth Benedict, Patterns of Culture, Boston: Houghton Mifflin, 1934.
- F. Kenneth Berrien, "Methodological and Related Problems in Cross-Cultural Research," International Journal of Psychology, 2(1, 1967), 33-43.
- John W. Berry, "Independence and Conformity in Subsistence Level Societies," Journal of Personality and Social Psychology, 7, 1967, 415-418.
- R.O. Blood and Y.J. Takeshita, "Development of Cross-Cultural Equivalence of Measures of Marital Interaction for U.S.A. and Japan," Transactions of the Fifth World Congress of Sociology, 4(1964), 333-334.
- Richard W. Brislin, Walter J. Lonner, and Robert M Thorndike, Cross-Cultural Research Methods, New York: Wiley, 1973.
- Lee J. Cronbach, Essentials of Psychological Testing, New York: Harper & Row, 1964.
- Andrew R. Davidson, "The Etic-Emic Dilemma: Can Methodology provide a Solution in the Absence of Theory?" in: Y.H. Poortinga (Ed.), Basic Problems in Cross-Cultural Psychology, Amsterdam: Swets & Zeitlinger, 1977, 49-54.
- Adriaan D. de Groot, Methodology: Foundations of Inference and Research in the Behavioral Sciences, The Hague: Mouton, 1969.
- George A. de Vos and Arthur A Hippler, "Cultural Psychology: Comparative Studies of Human Behavior," in G. Lindzey and E. Aronson (Eds.), The Handbook of Social Psychology, Reading: Addison-Wesley, 1969, Volume 4, 323-417.
- Susan P. Douglas, "Do Working Wives read different Magazines from Non-Working Wives?" Journal of Advertising, 6(Winter, 1977), 40-43 and 48.
- Bernard Dubois, "A Cultural Approach to the Study of Diffusion and Adoption of Innovations," in M. Venkatesan (Ed.), Proceedings of the 3rd Annual Conference of the Association for Consumer Research, 1972, 840-850.
- Joseph W. Elder, "Comparative Cross-National Methodology," Annual Review of Sociology, 2(1976), 209-230.
- Friedrich Engels, "The Origin of the Family, Private Property, and the State," in: Karl Marx and Friedrich Engels: Selected Works, 2(1949), Moscow: Foreign Languages Publishing House, 155-296.
- Claude Faucheux, "Cross-Cultural Research in Experimental Social Psychology," European Journal of Social Psychology, 6(3, 1976), 269-322.
- Nico Frijda and Gustav Jahoda, "On the Scope and Methods of Cross-Cultural Research," International Journal of Psychology, 1(2, 1966), 109-127.
- Robert T. Green and Philip D. White, "Methodological Considerations in Cross-National Consumer Research," Journal of International Business Studies, 7(Fall-Winter, 1976), 81-87.
- Donald J. Hempel, "Family Buying Decisions: A Cross-Cultural Perspective," Journal of Marketing Research, 11(August, 1974), 295-302.
- D. Horton, "The Functions of Alcohol in Primitive Societies: A Cross-Cultural Study," Quarterly Journal on the Study of Alcoholism, 4(1943), 199-220.
- F.L.K. Hsu, Clan, Caste and Club, Princeton: Van Nostrand. 1963.
- A. Inkeles, "Continuity and Change in the Interaction of the Personal and the Socio-Cultural Systems," in: B. Barber and A. Inkeles (Eds.), Stability and Social Change, Boston: Little Brown, 1971.
- George Katona, Burkhard Strumpel, and Ernest Zahn, Aspirations and Affluence, Comparative Studies in the United States and Western Europe, New York: McGraw-Hill, 1971.
- Kurt Koffka, Principles of Gestalt Psychology, New York: Harcourt Brace, 1935.
- Martin L. Maehr, "Culture and Achievement Motivation," American Psychologist, 29(December, 1974), 887-896.
- Karl Marx, Grundrisse: Foundations of the Critique of Political Economy, New York: Random House, 1973.
- C.A.O. Nieuwenhuijze, Cross-Cultural Studies, The Hague: Mouton, 1963, Chapter 8.
- Charles E. Osgood, "On the Strategy of Cross-National Research into Subjective Culture," Social Science Information, 6(1, 1967), 5-37.
- Charles E. Osgood, William H. May, and Murray S. Miron, Cross-Cultural Universals of Affective Meaning, Urbana, IL: University of Illinois Press, 1975.

Talcott Parsons, Societies: Evolutionary and Comparative Perspectives, Englewood Cliffs: Prentice-Hall, 1966.

L. Pearlin and M. Kohn, "Social Class, Occupation, and Parental Values: A Cross-National Study," American Sociological Review, 31(1966), 466-479.

Joseph T. Plummer, "Consumer Focus in Cross-National Research," Journal of Advertising, 6(Spring, 1977), 5-15.

Ype H. Poortinga, "Some Implications of three Different Approaches to Intercultural Comparison," in: J.W. Berry and W.J. Lonner (Eds.), Applied Cross-Cultural Psychology, Amsterdam: Swets & Zeitlinger, 1975, 327-332.

Karl R. Popper, The Logic of Scientific Discovery, New York: Dover, 1959.

Adam Przeworski and Henry Teune, "Equivalence in Cross-National Research," Public Opinion Quarterly, 30(4, 1966/67), 551-568.

S. Prakash Sethi, "Comparative Cluster Analysis for World Markets," Journal of Marketing Research, 8 (August, 1971), 348-354.

Jagdish N. Sheth, "A Review of Buyer Behavior," Management Science, 13(August, 1967), B-718-756.

Jagdish N. Sheth and S. Prakash Sethi, "An Approach to Developing a Theory of Cross-Cultural Buyer Behavior," Paper presented at the Annual Meeting of the American Psychological Association, Montreal, August 27-31, 1973.

Murray A. Straus, "Phenomenal Identity and Conceptual Equivalence of Measurement in Cross-National Comparative Research," Journal of Marriage and the Family, 31(May, 1969), 233-241.

Hans B. Thorelli, Helmut Becker, and Jack Engledow, The Information-Seekers -- An International Study of Consumer Information and Advertising Image, Cambridge, Mass; Ballinger Press, 1975.

Harry C. Triandis, Roy S. Malpass and Andrew R. Davidson "Psychology and Culture," Annual Review of Psychology, 24(1973), 355-378.

W. Fred van Raaij and Kassaye Wandwossen, "Motivation-Need Theories and Consumer Behavior," in H. K. Hunt (Ed.) Advances in Consumer Research, Vol. 5 Association for Consumer Research, 1978.

Max Weber, The Protestant Ethic and the Spirit of Capitalism, New York: Scribner's, 1958.

B.L. Whorf, "A Linguistic Consideration of Thinking in Primitive Communities," in: J.B. Carroll (Ed.) Language, Thought, and Reality: Selected Writing of Benjamin Lee Whorf, Cambridge, Mass: MIT Press, 1956.

Gerald Zaltman Christian R.A. Pinson, and Reinhard Angelmar, Metatheory and Consumer Research, Hinsdale, IL Dryden Press, 1973.

RESEARCH FOR CONSUMER POLICY FORMULATION:
AN EXPLORATION OF DIMENSIONS OF HOUSEHOLD ECONOMIC MANAGEMENT

Johan Arndt, University of Missouri-St. Louis
Oddvar Holmer, Norwegian School of Economics and Business Administration

Abstract

This study addresses a fundamental aspect of consumer behavior: administrative routines in the management of the economic affairs of the household. Several dimensions of economic management were uncovered, but the linkage of these dimensions to antecedent variables was weak.

Introduction

To date, most research on consumer behavior has been marketing management oriented and concerned with pre-purchase decision processes for brands. The relevant input to consumer policy formulation, however, is research with a more explicit "consumer economics" orientation. The study to be reported here is an example of such research, as it focuses on the fundamental problems in the administration of the economic affairs of the household. The study represents the pilot phase of a more comprehensive project.

The Household As A Small Business Firm

The conceptual framework underlying this study implies viewing the economic affairs of the household as a resource conversion process, in which the inputs are time, economic resources (income and capital), human capital, and service resources such as labor-saving appliances. The open-system organization of the household copes with the uncertainties and complexities of its market environment by developing information systems and by protecting its "technical core" with buffers of reserve money and reserve goods. To economize on time and effort, the household establishes, more or less formally and explicitly, administrative routines. These routines include rules for information acquisition, storage, and retrieval, and control procedures such as account keeping, budgeting, and advance planning, as well as mechanisms for conflict resolution. An important aspect of administrative routines is the organization of the household, particularly the delineation of authority and responsibility among its members for planning and implementation of purchases, contacts with finance institutions, etc.

The efficiency of the resource conversion process depends on the ability and willingness of the members of the household to plan and organize. And, in turn, this ability and willingness is believed to be related to background characteristics or antecedent variables of the household.

Method

The sample was not selected by any direct probability mechanism, but chosen in the following way: Four districts within the metropolitan area of Bergen, Norway were selected on a judgmental basis in order to ensure sufficient heterogeneity of the sample. In each district, a starting point was determined as a basis for selecting some 70 dwelling units by means of systematic sampling. Each household in the units selected received a structured questionnaire, which was collected on the following day. Of the 273 households contacted, 222 returned completed or partially completed questionnaires.

In the questionnaire, measurements were obtained for a set of 11 antecedent variables believed to be determinants of administrative routines. These variables included employment status of wife, age, household income, etc.

There were a total of 93 questionnaire items relating to administrative routines, categorized into three groups: convenience goods shopping, appliance purchasing, and items relating to financial management.

Findings

In this case it was desirable to search for the natural groupings occurring in the data rather than to rely on an a priori classification scheme. Therefore, factor analysis was used to track patterns of relationships and to reduce the data into basic dimensions of administrative procedures. The program used the VARIMAX in the SPSS package, extracted 32 factors, accounting for 73 per cent of the variance in the 93 items. However, as the contribution to explained variance dropped discontinuously after factor 7, only the first seven factors are included in the analysis.

These seven factors, which explained 36 per cent of the variance, were next related to the set of 11 antecedent variables using the factor scores for each factor as the dependent variables. This analysis was not successful, as the average proportion of the variance explained (in terms of the adjusted R^2 's) for the seven regressions was only .06. In the discussion of the factors below, antecedent variables not showing Beta coefficients significant at the 1 per cent level will not be discussed.

Factor 1 was clearly concerned with use of word of mouth information in appliance buying. High scores on this factor were positively related to the number of persons in the household and to the level of education, and both are possible indicators of a need for planning and the ability to plan.

Factor 2 covered cases in which neither husband nor wife was involved in administrative procedures.

The items loading highly on factor 3 were concerned with the use of shopping lists as a planning instrument. The factor was positively related to the number of persons in the household and to education.

The items showing the highest loadings on factor 4 related to spousal role differentiation in the financial management of the household.

The key dimension underlying factor 5 was the financing of major purchases "liberally", i.e. through instalment or bank credit rather than through saving or reducing other expenditures. Households living in the cheaper dwelling unit types such as small apartments were more inclined to use instalment credit, confirming that it is the economically less privileged consumer groups who tend to rely on credit. The negative correlation between scores on factor 5 and income adds further support for this interpretation.

Factor 6 seemed to reflect a "Make or buy" dimension, while the items loading highly on factor 7 related to the regular use of price deals or bargains in convenience goods shopping. The bargain hunters had smaller fixed expenditures per month (and were hence perhaps more likely to have more money at their disposal for bargains) and larger families.

Concluding Comments

The findings of an exploratory study such as the present one would be a shaky foundation on which to build sweeping generalizations. Instead, the results should be treated as hypotheses for future more rigorous study.

SESSION SUMMARY: CROSS-CULTURAL LIFE-STYLE RESEARCH

M. Venkatesan, University of Oregon

The Session Chairman, W. Fred Van Raaij, Tilburg University, pointed out that "cross-cultural" consumer behavior research is distinguished from "cross-national" research in the literature, but that he preferred the term "cross-cultural research," as the research not only can involve consumer behavior in different cultures but involves comparison of consumer behavior within subcultures. He pointed out that functional equivalence of constructs and instruments has to be established for real cross-cultural comparisons. The chairman pointed out other problems particularly relevant to cross-cultural consumer behavior research, viz., culturally ipsatized measurement problems, problems of operationalizations, "ethnocentrism" and the like.

suggesting possible uses of our findings for pragmatic business interests — "suffice at the present time that all involvement is indicated by our intellectual curiosity."

Johan Arndt, Norwegian School of Business Administration and Economics, focused on the ethnocentric bias manifested by the predominant marketing orientation to the study of cross-cultural consumer behavior because of almost all empirical studies have been conducted by U.S. academician or U.S. trained academicians. He emphasized, however, that properly conducted cross-cultural consumer behavior studies make it possible to validate findings obtained in one culture by replications in other cultures. The most important problem in Arndt's judgment was that of the unit of analysis in cross-cultural consumer behavior studies. While nuclear family may be an important unit of analysis in U.S. and some other cultures, the extended family unit may be the appropriate unit of analysis in many cultures.

Professor Eric Langeard, University of Aix-en-Provence pointed out that fairly similar types of situations can be studied with ease, such as opinion leadership for example, but he cautioned as the determinants of those situations might well vary across cultures even if the phenomenon is found to exist in consumer behavior contexts.

Professor G. van Veldhoven of Tilburg University focused on the problem of cross-cultural life style research. He pointed out that there are "conceptualization problems, and questions regarding "units of measurement." He illustrated such problems from his current study of life styles in Tilburg vs. Luill/Lille.

The discussant, Professor Venkatesan of the University of Oregon summed up the difficulties in conducting cross-cultural research. The main difficulties are: (1) assumptions of means-ends relationships; (2) problems of sample representativeness; (3) problems of validity and reliability of culturally ipsatized measurement, and (4) delineating micro and macro aspects of consumptive behavior. He cautioned against the use of instruments generated in the U.S. for purposes far removed from cross-cultural consumer behavior (e.g., AIDs generated in U.S. consumer studies). He also assailed the tendencies of researchers to "replicate" studies that were conducted in the U.S. without regard to the settings in other cultures, or without changing the instruments or generating new ones to suit the particular needs of these studies. However, he emphasized that cross-cultural consumer behavior studies can be conducted to test whether "procedures" utilized in U.S. consumer behavior studies work in other settings (cultures) and whether similar "manipulation" also work in different cultural settings. He concluded by pointing out that we are many many years away from

COMMENTS ON CROSS-CULTURAL CONSUMER RESEARCH

Johan Arndt, University of Missouri - St. Louis

Cross-cultural studies in consumer behavior are still in the early infancy stage. There is a paucity of published empirical studies; and worse, the few known studies suffer from methodological limitations. This not too encouraging state of the art is apparent in van Raaij's excellent comprehensive review in his "Cross-Cultural Research Methodology as a Case of Construct Validity."

Almost all empirical and theoretical consumer studies have been conducted in the U.S. by American or at least by U.S. trained academicians. This ethnocentric bias is manifested by the predominant marketing manager orientation and the focus on middle class consumers' problems in the mass-consumption society.

This occasion gives me an opportunity to reflect on the potentials and problems of cross-cultural consumer studies. The potential contributions should be self-evident:

1. Only cross-cultural studies may isolate the role of culture, a variable of interest in its own right as a determinant of consumer behavior.
2. Cross-cultural studies permit a more rigorous study of the role of various determinants of behavior by expanding the variation in them.
3. Cross-cultural studies permit extracting universals or trans-cultural concepts from idiosyncratic phenomena which are only the product of some unique historical or environmental factors.

The problems relating to cross-cultural research on consumer behavior may be divided into: (1) theoretical and conceptual problems concerning what to study and why; and (2) methodological problems relating to how to study the problems. As van Raaij has covered the methodological problems, particularly the problems of functional equivalence of samples and measurement instruments, in so much depth, the comments below will concentrate on the more fundamental problems of approach.

In the late 1940's and early 1950's, American business executives and academicians looked at the rest of the world mostly as a mission territory ready to accept the principles of modern business management and human relations. There is a danger that such "conceptual chauvinism" may repeat itself. Much current American consumer research is influenced by the perspective of the Marketing Concept, with its one-way control process conceptualization, and emphasis on prepurchase decision processes for brands and reliance on concepts such as "brand loyalty", "information overload", and "consumer satisfaction". To a Eastern European car buyer having just one brand to choose and little information, and to a poor Southeast Asian farmer such concepts may be meaningless. While in the U.S., much consumer behavior is an individual or nuclear family affair, the extended family may be the appropriate unit in other cultures.

Many cross-cultural studies give the impression of being opportunistic, a result of short-term sabbaticals abroad or sudden access to some secondary international data. Second, the research is reductionistic in the sense that the consumer behavior studied is not explicitly related to aspects of the consumers' resource situation and their market environment. Hence it difficult, if not impossible, to interpret the differences found.

In the future, perhaps the only feasible approach in cross-cultural consumer research, is the launching of larger scale projects operated by cross-national research teams. To explore the interrelations between political factors, the socio-economic environment, marketing institutions, and consumer behavior, one possible solution is more emphasis on detailed comparative case studies of selected marketing systems.

INFERENTIAL BELIEF FORMATION IN THE CUE UTILIZATION PROCESS

Jerry C. Olson, Pennsylvania State University¹

Abstract

The extant research on cue utilization is briefly reviewed. Generally, this research has ignored inferential processes by which beliefs are formed about concepts not present in the immediate environment. A model of the cue utilization process is presented which explicitly includes the inferential belief formation process. Memory schemata are proposed as the basic memory structure which enables inferences or attributions to be made. A discussion of how research on the inference process could be conducted follows, and the paper concludes by identifying several research issues that could be examined from the conceptual perspective provided by the model.

Introduction

How do people acquire and use information about their environment to make various types of judgement choices? What factors influence these processes? Issues regarding the cue utilization process have fascinated psychologists for many years. More recently, consumer researchers have become quite interested in how consumers use informational cues to accomplish various tasks. This paper briefly reviews the types of cue utilization studies and summarizes the typical approach taken. It is then suggested that most of the cue utilization research has ignored the potentially important process of inferential belief formation. A flow-chart model is presented which illustrates how inferential processes might be included in the cue utilization process. The paper concludes by briefly discussing several research topics which could be investigated from the perspective provided by the proffered model.

Overview of Cue Utilization Research

Define somewhat loosely, the cue utilization process may be considered to encompass the cognitive processes involved in obtaining information from the external environment and in using the information (in a cognitive, information processing sense) to produce a particular behavior, e.g., an evaluation or a choice. Defined at this broad general level, the cue utilization process is involved in virtually all information processing phenomena.

Given this broad perspective, then, a wide variety of research areas can be seen as focussing on aspects of the cue utilization process. Included are research areas such as impression formation (Asch, 1946; Wyer, 1974; Zanna and Hamilton, 1977), decision theory and judgment formation (Slovic, Fischhoff, and Lichtenstein, 1977; Kahneman and Tversky, 1973), integration theory (Anderson, 1971), attribution theory (Frieze and Weiner, 1971; Fishbein and Ajzen, 1975), syllogistic relationships among beliefs (Henninger and Wyer, 1975; McGuire, 1960), or social inference (Gollob, Rossman and Abelson, 1973; Wyer, 1975). Although there appears to be relatively little cross-communication between these fields, each research area deals, in the broad sense suggested above, with how people use informational cues. That is, the investigator in each research area is interested in how

people select and comprehend environmental information and possibly use the information in forming the response requested (usually an evaluative judgment or overt choice behavior).

In the consumer research literature, cue utilization research is not quite so compartmentalized, although studies in virtually all of the above-mentioned research traditions have been published. Essentially, just three major research areas have evolved. The largest part of consumer-oriented research is focused on informational effects on consumers' product judgments, usually of quality (see Olson, 1972), and particularly on the effects of price cues on such judgments (Monroe, 1973; Olson, 1977). Another rapidly developing research area involves study of consumer information acquisition behavior by Jacoby and his colleagues (cf. Jacoby, Chestnut, Weigl, and Fisher, 1976). A third research area which involves cue utilization processes seeks to identify the combinatorial "rules" used by consumers to integrate or combine various items of information in order to accomplish some task (cf. Bettman, 1970; Wright, 1975; Scott and Wright, 1976; Sheth and Raju, 1974).

Despite the specific focus of most of this research, it is clear that the cue utilization process is involved in virtually all behavioral phenomena. That is, nearly any behavior will require processing of informational cues selected from the external environment. Thus, the process of cue utilization is of broad importance. However, for purposes of this paper, our attention is restricted to those controlled situations in which the subject is given a particular task to accomplish that requires cue processing, as in an experimental study of cue utilization behavior.

Summary of Cue Utilization Research

Much of the cue utilization research in marketing must be characterized as simple. Many studies involved presenting consumer subjects with task environments containing only one relevant cue (e.g., Leavitt, 1954; Peterson, 1970). More recent quality perception studies have tended to provide more cue stimuli and, thus, have created somewhat more complex and ecologically valid task environments (see Olson, 1977). Consumers have been provided with two cues (Enis and Stafford, 1969; Valenzi and Andrews, 1971), three cues (Szybillo and Jacoby, 1974; Wheatly and Chiu, 1977) and four cues (Jacoby, Olson, and Haddock, 1971) in an environment with the task to evaluate the quality of one or several alternative brands. The research on cue weighting in the integration process has used task environments containing even greater numbers of cues, as many as 10 or 11 (see review by Scott and Wright, 1976). In most of these cue utilization studies, the consumer's task has been to give evaluative judgments of the product (e.g., menu appeal, Green, Wind, and Jain, 1972). Thus the typical dependent variable has traditionally been a rather simple evaluative scale such as good-bad, like-dislike, or high quality-low quality. Essentially then, many of these studies may also be considered as examining the attitude formation process.

¹Associate Professor of Marketing.

A Simple Model of the Cue Utilization Process

In sum, the cue utilization process can be considered to

involve two basic processes--(a) the acquisition of cues and (b) the integration of the information derived from the cues to form the desired response (Olson and Jacoby, 1972). Actually, however, this view is highly simplified and, in fact, tends to hide several complex information handling sub-processes that may occur upon exposure to stimulus cues in a task context. Current interest in the cognitive states and processes involved in information processing (cf. Jacoby and Olson, 1977) suggests that several hierarchically-arranged stages are involved. Figure 1 presents a simple model of some of the stages involved in cue utilization. This model should be considered only as a convenient illustrative scheme to communicate some of the complexities involved. Most of the extant research on cue utilization can be described within the confines of this model.

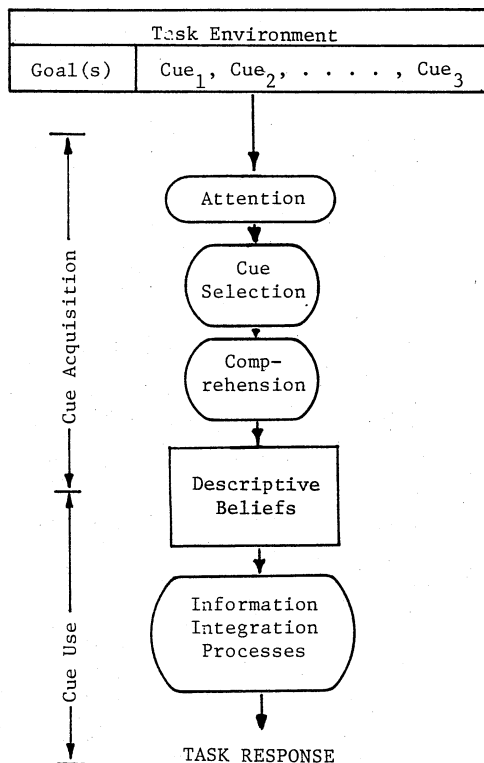


Figure 1. A SIMPLIFIED MODEL OF THE CUE UTILIZATION PROCESS

Briefly, presentation of a task goal (Which brand do you think is of highest quality?) and a task environment (four brands of shoes with price, brand name, and other informational cues), focuses the attention of the consumer/subject and initiates the goal-oriented cue utilization process. The initial sensory processing of the environment is probably extremely rapid, especially for familiar task environments, and may not be available to conscious awareness. Therefore, this stage typically goes unrecognized and unreported (Broadbent, 1977). One can think of this preconscious processing as orienting the organism to the general class of stimuli. This will be elaborated later in the paper.

Following, relatively more conscious selection of specific informational cues (dimensions, attributes, etc.) occurs.

These specific cues are subjected to further processing, the outcome of which is the comprehension of the selected cues, i.e., the assignment of categories of meaning to each cue. This meaning--now encoded information--may be stored as descriptive beliefs about the object. Descriptive beliefs encode the perceived association between the task concept and specific attributes or dimensions present in the task environment. These beliefs may then be integrated together or combined to accomplish the task, such as rating the quality of each brand.

Several points about the cue utilization process are made obvious by Figure 1. First, the cue utilization process is probably highly selective of certain cues. Interestingly very few studies have examined factors which may influence the cue selection process (cf. Cox, 1962; Olson, 1972). Second, the outcome of the integration sub-process is highly dependent upon the precise encoded informational form of the selected cues, i.e., the specific beliefs about the stimulus concept (Olson, in press). That is, environmental cues may be transformed during encoding and assigned a meaning that may or may not be similar to the physical form of the cue. For instance, a price of 29¢ may be encoded as "cheap," rather than "29¢," and integrated with other product information in the "cheap" form (Jacoby and Olson, 1977). Third, certain complex informational cues may not be encodable in a meaningful way by a perceiver (e.g., nutritional cues, Jacoby, Chestnut and Silberman, 1977), and thus may have little effect on the task response. For the most part, these three points have not been clearly recognized nor explicitly considered in previous cue utilization research.

Inferential Belief Formation

A fourth implication, implicit in Figure 1, involves the formation of inferential beliefs, a process that has generally been ignored in cue utilization research. That is, it is possible that in processing the given cues from a task environment, one might develop beliefs about other aspects of the task stimuli not represented by environmental cues. This is the process of inferential belief formation (see Fishbein and Ajzen, 1975). The importance of this process is that both the descriptive beliefs (directly related to environmental cues) and the inferential beliefs (inferred from the environmental cues) will, according to the basic expectancy-value theory, both influence global evaluation of attitude toward the concept of interest (cf. Fishbein and Ajzen, 1975; Lutz, 1975).

Although as yet not explicitly studied in the consumer behavior literature, there have been several empirical demonstrations of inferential belief information. For instance, Mazis and Adkinson (1976) found that exposure to a corrective ad about a mouthwash's inability to block colds and sore throats also changed beliefs regarding the brand's ability to "kill germs." Informational cues regarding killing germs were not present in the task environment and thus the latter beliefs must have been formed via some inferential process. Lutz (1975) found that a persuasive message regarding a single attribute of detergent affected the strengths of other beliefs about the product. As a final example, Mitchell and Olson (1977) found that an ad containing only a symbolic visual image and the brand name created stronger beliefs about the symbolically-communicated product attribute than an explicit verbal claim about the same attribute. In this case, the beliefs were inferred from a pictorial cue. Fishbein and Ajzen (1976) describe many other studies from the social psychology literature that also demonstrate the formation of inferential beliefs.

In sum, it is clear that people may form beliefs about the task object/concept that are not represented by cues in the task environment. This phenomenon is termed inferential belief formation. The next task, then, is to develop

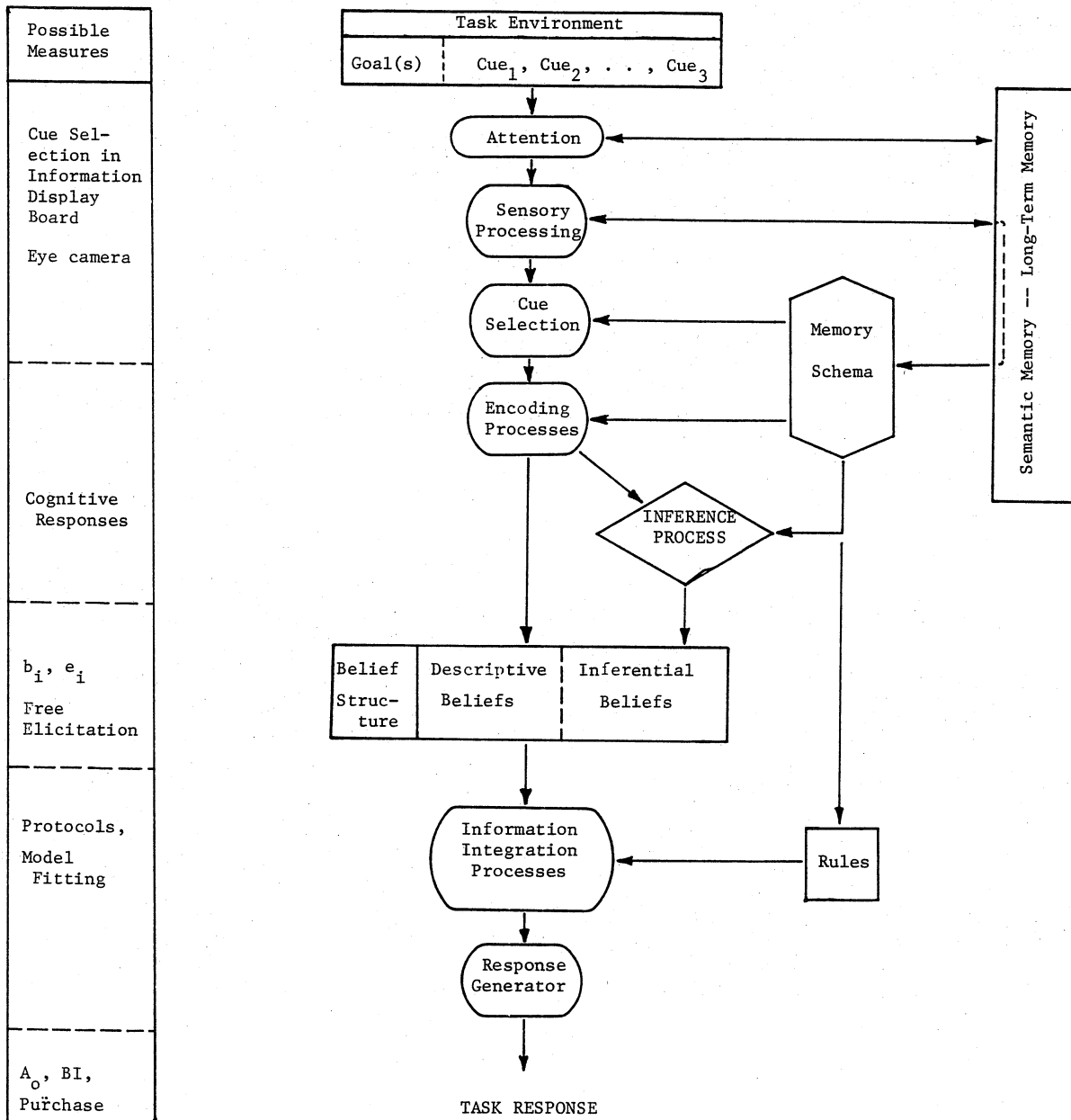


Figure 2. CONCEPTUAL MODEL OF THE CUE UTILIZATION PROCESS, INCLUDING THE INFERENCE BELIEF FORMATION PROCESS

a theoretically-based model of the cue utilization process which can logically account for such phenomena.

Inferential Belief Formation in the Cue Utilization Process

It is obvious that an explanation for inferential belief formation requires a consideration of previously learned information upon which one can logically base, or from which one can derive, inferences. That is, an adequate model of the cue utilization process must account for the influence of memory on the inferential belief formation process. Fortunately, the cognitive psychology field has developed numerous memory concepts and models. Several of these ideas appear useful in developing such a more theoretically complete model of cue utilization process--one that includes the possibility of inferential belief formation (cf. Olson, 1977).

Memory Schemata

Perhaps the most useful concept for the purpose of this paper is that of memory schemata. Norman and Bobrow (1975; Bobrow and Norman, 1975) developed this idea of highly organized substructures of knowledge stored in semantic memory. Schemata for products or brands would

²Tulving (1972) suggested a distinction between episodic memory (memory for past events in our lives) and semantic memory (stored knowledge about the world). It would seem that inferential processes about products are more likely to be influenced by semantic than episodic memory. This is particularly so given the perspective of focusing on specific product characteristics followed by most product perception or multi-attribute research. It would seem that semantic memory, in which is stored the meanings linked to the concept via its associated attribute/characteristics, is likely to influence the formation of inferential beliefs.

contain previously learned knowledge about the concept, plus, the interrelationships between these items of knowledge, stored in an organized logical framework. Interestingly, schemata may also contain "rules" for responding to the stimulus concept--e.g., a learned tendency to favorably evaluate a high price for a particular product. In this sense, schemata are "active" cognitive units which can influence behavior rather "directly," once activated. A schema may be activated simply by exposure to the stimulus representing to concept-context combination. Thus, schemata tend to be highly context specific. One might find different knowledge and different "rules" stored in the schemata for (a) toothpaste to use in the morning and (b) toothpaste to use before a social engagement.

Figure 2 represents an initial attempt to model how schemata may be integrated into the cue utilization process in a way that may explicate the inferential belief formation process. In Figure 2, the process begins with an experimental setting in which a subject is given a specific task (goal) and an informational cue environment within which to accomplish the task. Consider the previous example of evaluating the quality of four pairs of men's dress shoes given a set of cues about the shoes. The early attentive and sensory processing will activate (make available for conscious processing) the relevant schema, in this case the schema for men's dress shoes. This, of course, presumes that the subject has acquired a schema for men's dress shoes. If the consumer has not developed a special knowledge structure for dress shoes, the task environment might activate a more general schema for men's shoes.

Once activated, the schema "controls" the processes of cue selection and cue comprehension or encoding. For instance, if a relevant component of the schema for men's dress shoes is leather soles, a desired feature, then the consumer may "be directed by the schema" to select cues regarding the composition of the soles for the alternative brands. Once a cue is selected, the encoding/comprehension process also proceeds "in light of" the activated, hopefully relevant schema. For instance, a price of \$49.00 may be encoded as "about average" given one consumer's schema for dress shoes. Another consumer may interpret the same price as "excessively excessive"--due to a difference schema for dress shoes. Moreover, the same \$49.00 price might be encoded differently by the same consumer when associated with casual sports shoes than with dress shoes if a different schema has been activated. Clearly, how we encode incoming informational cues, i.e., the meaning we assign to cues, is a function of our past experiences (cf. Fishbein and Ajzen, 1975). The schemata concept can be seen as a theoretical representation of the cognitive structure created by those past experiences.

Now, as a result of selecting and encoding some or all of the informational cues in the task environment, beliefs are formed. Those directly related to environmental cues are termed descriptive beliefs to distinguish them from inferential beliefs. It may be that no inferential beliefs are formed. For example, in the case where one desires whole wheat bread, one may acquire and encode only those cues related to the grains used in manufacturing the bread and select the first whole wheat alternative found. Other beliefs about the breads may not be formed. However, in most situations it seems likely that inferential beliefs will be formed. In fact, for cases involving activation of well-developed schemata, the inferential process may proceed without much conscious analytical thinking, apparently automatically.

The inferential process itself, although still not completely explicated, can be likened to an attributional process in which incoming encoded information from the

task environment is compared/fitted into the established knowledge structure of the schema and inferences drawn about other concepts not present in the immediate environment.

Both descriptive beliefs and inferential beliefs (if any) form the cognitive content upon which the integration process operates. As noted above, the schema may also provide the combinatorial "rules" for integrating the information. The resultant cognitive state created by integrating the information is then "passed" to a "response generator process" which converts the cognitive state into the appropriate response called for by the experimental task--e.g., a check mark on a 7-point product quality scale.

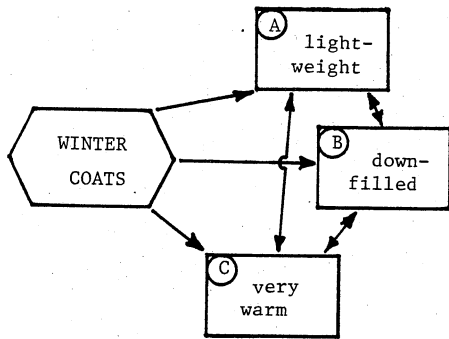
The Inference Process

Now that the stages represented in Figure 2 have been discussed, let us examine more closely how a memory schema may allow inferences to be made regarding product attributes not represented in the immediate task environment. Part a of Figure 3 illustrates a portion of a memory schema for the generic product category, winter coats. This schema contains a consumer's knowledge regarding coats that are insulated with goose down. This knowledge is stored in a highly organized framework represented by distinct associations between concepts. Thus there are linkages (logical associations, or beliefs) between the concept, down-filled, and the product attributes of weight and warmth. These concepts associated with winter coats and their interrelationships were previously acquired/learned via classical and/or instrumental learning principles (cf. Fishbein, 1967; Olson and Mitchell, 1975) as a function of one's past experiences. In one sense, these associations between the knowledge components of the schema are very much like general product expectations (Olson and Dover, 1976). Now, suppose that this particular subject possessing this schema is exposed to a task environment and instructed to evaluate a particular winter coat. Assume, moreover, that the only cue provided by the environment is that the coat is down-filled. Upon processing (acquiring and encoding) this cue, the subject will form a belief, probably a strong belief, that this brand does indeed possess the attribute, "down-filled" (see Part b of Figure 3). In addition, exposure to the task environment will probably activate the previously-learned schema of winter coats illustrated in Part a of Figure 3. When the descriptive belief that the coat is down-filled is "fitted into" the schema, the previously learned relationships "enable" the consumer to infer certain other qualities of the coat. This inferential belief process is represented in Part c of Figure 3, which illustrates that the two beliefs (represented as subjective probabilities) P_{XB} and P_{BA} , about the associations between concept X and B and B and A , allow an "inference" to be made about the association between X and A , P_{XA} . This latter probability is the inferential belief about a product attribute not explicitly present in the task environment.

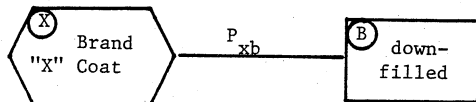
In other disciplines, notably social psychology, numerous studies have focused on peoples' ability to form inferences and of the factors which influence that ability. Some years ago, McGuire (1960) developed a probability model of the relationships between syllogistically-related beliefs. More recently, Wyer (cf. 1974) has extended this model (Wyer and Goldberg, 1970). Wyer has shown in several studies (cf. Wyer, 1975) that peoples' subjective estimates of some of the associations between concepts in a syllogism predict with considerable accuracy the strength of the other association in the syllogism.

A Probability Model of the Inference Process

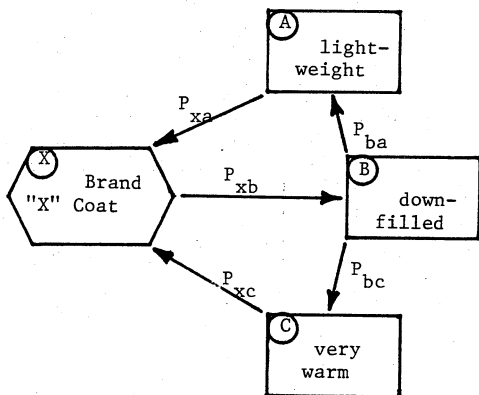
The McGuire-Wyer model appears useful in explicating how



a. MEMORY SCHEMA FOR GENERIC PRODUCT -- WINTER COATS



b. BELIEF FORMATION (P_{xb}) CAUSED BY PROCESSING CUE "B" FROM THE TASK ENVIRONMENT



c. BELIEF STRUCTURE (P_{xb} , P_{xa} , P_{xc}) CREATED THROUGH INFERENCE (arrows indicate theoretical causal flow).

Figure 3. THE INFLUENCE OF A MEMORY SCHEMA ON THE INFERENCE BELIEF FORMATION

a memory schemata can account for inferential belief formation. Figure 4 illustrates a schema containing three concept, A, B, and C. Also illustrated are some of the logical links or associations between the concepts, stated in probabilistic terms. For example, the subjective estimate of the likelihood of C given A ($P_{C/A}$) is .60. The likelihood of C in the absence of A ($P_{C/A'}$) is only .20. These subjective associations indicate that concept "C" is perceived to be somewhat contingent upon concept A. In contrast, Figure 4 indicates that concept B is not perceived as highly contingent upon the presence of A--note that $P_{B/A'} = .70$, only .10 less than $P_{B/A} = .80$. Further, suppose that exposure to an experimentally-controlled task environment creates a strong descriptive belief that concept X possesses concept A--e.g., $P_{XA} = .90$. The question of interest involves the belief strengths of the two inferentially-created beliefs, P_{XC} and P_{XB} .

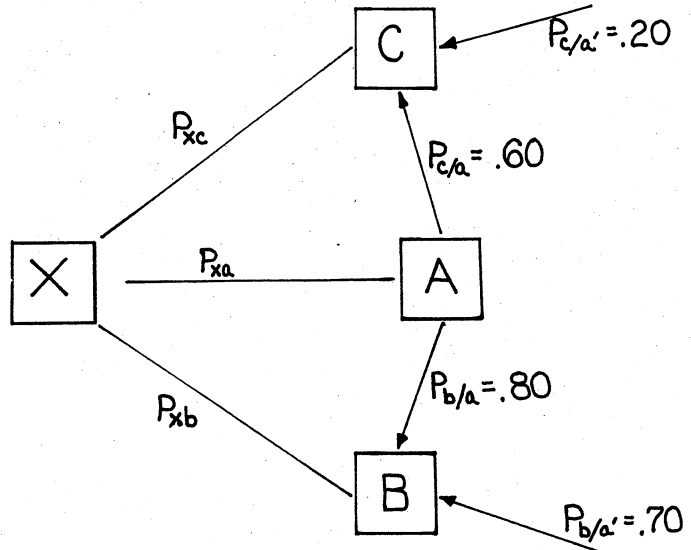


Figure 4. PROBABILISTIC RELATIONSHIPS BETWEEN THREE CONCEPTS STORED IN A MEMORY SCHEMA.

The McGuire-Wyer Subjective Probability Model of Cognitive Functioning (or Inferential Belief Formation), allows predictions of P_{XC} and P_{XB} . The model, based upon conditional probability theory; is as follows:

$$P_{xc} = P_{xa} \cdot P_{c/a} + (1 - P_{xa}) \cdot P_{ca'} \quad (1)$$

Verbally, the association strength between X and C is equal to the strength of association between X and A times the conditional probability of C given A, plus the probability of no link between X and A times the conditional probability of C given absence of A. This model produces the following prediction of the inferential belief, P_{xc} :

$$P_{xc} = (.90)(.60) + (.10)(.20)$$

$$P_{xc} = .56$$

Likewise, the inferentially-formed belief regarding the association between concepts X and B can be estimated by the same model:

$$P_{xb} = P_{xa} \cdot P_{b/a} + (1 - P_{xa}) \cdot P_{b/a'} \quad (2)$$

$$P_{xb} = (.90)(.80) + (.10)(.70)$$

$$P_{xb} = .79$$

Once estimated, these belief strengths can be compared with direct ratings of the perceived associations. In several studies, Wyer (cf. 1974) reported reasonably strong relationships between model predictions and independent ratings.

Using this probabilistic approach, one can study the organization of stored knowledge in schemata held in semantic memory, as well as the use of that stored information in the process of inferential belief formation. Briefly, the procedure could be as follows. First, using a relatively non-reactive method such as a free elicitation procedure ("say everything that comes to mind when I say _____"), identify the concepts stored in semantic memory that are associated with the concept. Then, ask consumers to subjectively rate the strengths of associa-

tion (perhaps in both directions) between pairs of concepts. These procedures will yield a probabilistic view of the content and structure of the knowledge stored within a memory schema for a particular object concept. With this data, one should be able to predict the inferential beliefs and the strengths of those beliefs created by an experimental message directed at one of the target concepts within the schema. Certainly, this type of approach should be examined in future research.

Other Research Issues

Product Familiarity

The model proposed in Figure 2 has implications for phenomena other than inferential belief formations. For example, the concept of "product familiarity" has frequently been used by consumer researchers as an "explanation" of some observed data or as a blocking variable in experimental designs. "Product expertise" is another concept commonly used in similar ways. Seldom, however, is either concept precisely defined in clear conceptual terms, and virtually never in terms of the cognitive state associated with each variable.

However, the notion of memory schemata, which possess varying degrees of complexity in terms of amounts of stored knowledge and "interrelatedness" of the stored information, may be used to represent the cognitive state created by degrees of "familiarity" or "expertise" with a product category. It should be noted that the relationship between amount of past experience and complexity of memory schemata is not clear. It may be that familiar, expert customers have in fact, less complex cognitive structures (schemata) than moderately familiar consumers (cf. Hayes-Roth, 1977; Olson and Dover, 1978). Although the degree of complexity is in doubt, it is essentially an empirical question that appears determinable with relative ease. In any case, the schemata of the expert, whatever their complexity, should lead to differential cue utilization behavior compared to a non-expert, and these differences should be predictable once the schemata are measured. This approach to issues of product familiarity/expertise has promise of providing deeper levels of understanding of these important concepts.

Development of Schemata

A variety of important related issues involve the development of memory schemata. We might ask, "What types of product information and experiences lead to the acquisition of well-developed, complex schemata that are useful in reacting to a product in varying circumstances and settings? How much overt, conscious, cognitive analysis is necessary in schemata development? Do low-involvement products have simpler schemata? Does "low involvement processing" lead to non-complex schemata? These questions focus on the fascinating issues involved with the processes of schemata development (cf. Olson, in press) and certainly warrant future research attention.

Activation of Schemata

A final issue to be briefly discussed involves the activation of a schema for conscious processing. A critical question to be answered is, "What cues trigger or activate a schema?" Are purchase context cues or broad purchase goals critical factors in determining the precise schema that is activated? For instance is "buying for self" vs. "buying for a gift" a distinct cue to activating different schemata for products such as ball-point pens or table wine? Or, is the price level a broad cue for activating a particular schemata for certain product categories such as automobiles or men's shirts. Such questions (and many others unstated) are of high rele-

vance to marketers and must be addressed in future research before the schemata concept can be widely useful.

Summary

This paper presents several ideas intended to broaden the approach taken in cue utilization research. It was suggested that inference processes may be involved in the cue utilization process which may create beliefs about concepts not present in the task environment. A model was presented which positioned the inferential belief formation process within the cue utilization process. The concept of memory schemata was proposed as a basis for a viable theoretical explanation for the inference process. The issues raised by this perspective are of interest to consumer researchers for both theoretical and applied reasons. This paper will serve its intended purpose if others are stimulated to study the inferential belief formation process, its effects, and the factors which affect it.

References

- Icek Ajzen and Martin Fishbein. "A Bayesian Analysis of Attribution Processes," Psychological Bulletin, 82 (1975), 261-77.
- Norman H. Anderson. "Integration Theory and Attitude Change," Psychological Review, 78 (1971), 171-206.
- Solomon E. Asch. "Forming Impressions of Personality," Journal of Abnormal and Social Psychology, 41(1946), 258-90.
- James R. Bettman. "Information Processing Models of Consumer Behavior," Journal of Marketing Research, 7 (August 1970), 370-6.
- Daniel G. Bobrow and Donald A. Norman. "Some Principles of Memory Schemata," in D. G. Bobrow and A. M. Collins, eds., Representation and Understanding: Studies in Cognitive Science (New York: Academic Press, 1975).
- Donald E. Broadbent. "The Hidden Preattentive Processes," American Psychologist, (1977), 109-18
- Ben M. Enis and James E. Stafford. "Consumer's Perception of Product Quality as a Function of Various Informational Inputs," in P. R. McDonald (ed.), Marketing Involvement in Society and the Economy (Chicago: American Marketing Association, 1969), 340-4.
- Martin Fishbein and Icek Ajzen. Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research, (New York: Addison-Wesley, 1975).
- Irene H. Frieze and B. Weiner. "Cue Utilization and Attributional Judgments for Success and Failure," Journal of Personality, 39(1971), 591-605.
- H. F. Gollub, B. B. Rossman and Robert P. Abelson. "Social Influence as a Function of the Number of Instances and Consistency of Information Presented," Journal of Personality and Social Psychology, 27 (1973), 19-33.
- Paul E. Green, Yoram Wind and Arun Jain. "Preference Measurement of Item Collections," Journal of Marketing Research, 9(1972), 371-7.
- Barbara Hayes-Roth. "Evolution of Cognitive Structures and Processes," Psychological Review, 84 (May 1977), 260-78.

- Marilyn Henninger and Robert S. Wyer, Jr. "The Recognition and Elimination of Inconsistencies Among Syllogistically Related Beliefs: Some New Light on the 'Socratic Effect'." Journal of Personality and Social Psychology, 54(1976), 680-93.
- James L. Jaccard and Martin Fishbein. "Inferential Beliefs and Order Effects in Personality Impression Formation." Journal of Personality and Social Psychology, 31(December 1975), 1031-40.
- Jacob Jacoby, Robert W. Chestnut and William Silberman. "Consumer Use and Comprehension of Nutritional Information." Journal of Consumer Research, 4(1977), 119-28.
- Jacob Jacoby, Robert W. Chestnut, Karl C. Weigl, and William Fisher. "Pre-purchase Information Acquisition: Description of a Process Methodology, Research Paradigm, and Pilot Investigation, in Beverlee B. Anderson, ed., Advances in Consumer Research: Vol. III (Cincinnati: Association for Consumer Research, 1976), 306-14.
- Jacob Jacoby and Jerry C. Olson. "Consumer Reaction to Price: An Attitudinal, Information Processing Perspective," in Yoram Wind and Marshall Greenberg, eds., Moving A Head with Attitude Research, (Chicago: American Marketing Association, 1977), 73-86.
- Jacob Jacoby, Jerry C. Olson and Rafael A. Haddock. "Price, Brand Name, and Product Composition Characteristics as Determinants of Perceived Quality," Journal of Applied Psychology, 55(December 1971), 570-9.
- Daniel Kahneman and Amos Tversky. "On the Psychology of Prediction," Psychological Review, 80(1973), 237-51.
- Harold J. Leavitt. "A Note on Some Experimental Findings about the Meaning of Price," Journal of Business, 27(1954), 205-10.
- Richard J. Lutz. "First-Order and Second-Order Cognitive Effects in Attitude Change," Communication Research, 2(July 1975), 289-99. (b)
- Michael B. Mazis and Janice E. Adkinson. "An Experimental Evaluation of a Proposed Corrective Advertising Remedy," Journal of Marketing Research, 13(1976), 178-83.
- William J. McGuire. "A Syllogistic Analysis of Cognitive Relationships," in Carl J. Hovland and Milton J. Rosenberg, eds., Attitude Organization and Change (New Haven, Conn.: Yale University Press, 1960).
- Andrew A. Mitchell and Jerry C. Olson. "Cognitive Effects of Advertising Repetition," in William D. Perrault, Jr., ed., Advances in Consumer Research: Vol. IV. (Atlanta: Association for Consumer Research, 1977), 213-20.
- Kent B. Monroe. "Buyer's Subjective Perceptions of Price," Journal of Marketing Research, 10(1973), 70-80.
- Donald A. Norman and Daniel G. Bobrow. "On the Role of Active Memory Processes in Perception and Cognition," in Charles N. Cofer, ed., The Structure of Human Memory. (San Francisco: W. H. Freeman, 1975).
- Jerry C. Olson. "Cue Utilization in the Quality Perception Process: A Cognitive Model and an Empirical Test." Unpublished doctoral dissertation, Purdue University, 1972.
- Jerry C. Olson. "Price as an Informational Cue: Effects on Product Evaluations," in Arch G. Woodside, Jagdish N. Sheth, and Peter D. Bennett, eds., Consumer and Industrial Buying Behavior. (New York: Elsevier North-Holland, 1977).
- Jerry C. Olson. "Theories of Information Encoding and Storage: Implications for Consumer Research," in A. A. Mitchell (ed.), Effects of Information of Consumer and Market Behavior. (Chicago: American Marketing Association, in press).
- Jerry C. Olson and Philip A. Dover. "The Effects of Expectation Creation and Disconfirmation on Belief Elements of Cognitive Structure." In B. B. Anderson, ed., Advances in Consumer Research: Vol. III. (Cincinnati, Ohio: The Association for Consumer Research, 1976), 168-175.
- Jerry C. Olson and Philip A. Dover. "Attitude Maturation: Changes in Underlying Belief Structures Over Time," in H. Keith Hunt, ed., Advances in Consumer Research: Vol V. (Association for Consumer Research, 1978).
- Jerry C. Olson and Jacob Jacoby. "Cue Utilization in the Quality Perception Process," in M. Venkatesan, ed., Proceedings, Third Annual Conference. (Iowa City, Iowa: Association for Consumer Research, 1972), 167-79.
- Jerry C. Olson and Andrew A. Mitchell. "The Process of Attitude Acquisition: The Value of a Developmental Approach to Consumer Attitude Research, in Mary J. Schlinger, ed., Advances in Consumer Research: Vol. II. (Chicago: Association for Consumer Research, 1975), 249-264.
- Robert A. Peterson. "The Price-Perceived Quality Relationship: Experimental Evidence," Journal of Marketing Research, 7(1970), 525-8.
- Jerome E. Scott and Peter Wright. "Modeling an Organizational Buyer's Product Evaluation Strategy: Validity and Procedural Considerations." Journal of Marketing Research, 13(1976), 211-24.
- Jaddish N. Sheth and P. S. Raju. "Sequential and Cyclical Nature of Information Processing Models in Repetitive Choice Behavior," in S. Ward and P. Wright (eds.), Advances in Consumer Research: Vol. I. (Urbana, Illinois: Association for Consumer Research, 1974), 348-58.
- Paul Slovic, Bernard Fishchhoff and Sara Lichtenstein. "Behavioral Decision Theory," Annual Review of Psychology 28(1977), 1-39.
- George J. Szybillo and Jacob Jacoby. "Intrinsic vs. Extrinsic Cues as Determinants of Perceived Product Quality," Journal of Applied Psychology, 59(1974), 74-8.
- Endel Tulving. "Episodic and Semantic Memory," in Endel Tulving and W. Donaldson, eds., Organization of Memory. (New York: Academic Press, 1972).
- Enzo R. Jalenzi and I. R. Andrews. "Effect of Price Information on Product Quality Ratings," Journal of Applied Psychology, 55(1971), 87-91.
- John J. Wheatley and John S. Y. Chiu. "The Effects of Price, Store Image, and Product and Respondent Characteristics on Perceptions of Quality," Journal of Marketing Research, 14(1977), 181-6.
- Peter Wright. "Consumer Choice Strategies: Simplifying vs. Optimizing," Journal of Marketing Research, 12(February 1975), 60-7. (a)
- Robert S. Wyer, Jr. Cognitive Organization and Change: An Information Processing Approach. (Potomac, Md.: Lawrence Erlbaum, 1974).

Robert S. Wyer, Jr. "Some Informational Determinants of One's Own Liking for a Person and Beliefs That Others Will Like This Person." Journal of Personality and Social Psychology, 31(1975), 1041-53.

Robert S. Wyer, Jr. and L. Goldbert. "A Probabilistic Analysis of the Relationships Among Beliefs and Attitudes," Psychological Review, 77(1970), 100-20.

Mark P. Zanna and David L. Hamilton. "Further Evidence for Meaning Change in Impression Formation," Journal of Experimental Social Psychology, 13(1977), 224-38.

SELF-PERCEPTION PROCESSES IN CONSUMER BEHAVIOR:
INTERPRETING ONE'S OWN EXPERIENCES

Carol A. Scott, The Ohio State University

Abstract

Although consumer behavior has been conceptualized as dynamic in nature, little empirical work has been done on the process by which behavior in one time period forms the basis for subsequent attitudes and actions. Recent developments in self-perception theory, however, provide the analytical tools necessary for investigating this process. The basic postulates of and evidence for self-perception are reviewed, and current research issues are identified and discussed in this paper. Implications for consumer research are then considered.

Introduction

Consumer behavior has been conceptualized traditionally as a dynamic process in which actions are affected by what is learned from previous behavior as well as by information from external sources (e.g., advertisements, friends, relatives, etc.). This property of dynamism is incorporated implicitly in most research (e.g., studies which control for past experience by using new or fictitious brand names), and explicitly in theoretical statements which include feedback loops from behavioral to cognitive variables (e.g., Engel, Kollat & Blackwell, 1973; Howard & Sketch, 1969; Nicosia, 1966) and in research utilizing Markov, linear learning, or stochastic models of brand choice (e.g., Bass, 1974; Kuehn, 1962). While this link between past and future behavior is acknowledged, little research attention has been focused on the process by which behavior in one time period forms the basis for subsequent attitudes and actions.

Recent theoretical developments provide the analytical tools necessary for exploring this rich and exciting area. Specifically, self-perception theory (Bem, 1965, 1972) attempts to explicate how an individual interprets his own behavior, how he assigns meaning to that behavior, and under what conditions the individual accepts experiential information as valid and worthy of incorporation into his attitude and behavior set. Self-perception theory's central province is the feedback loop, often postulated by theory and often neglected empirically.

Although self-perception theory has been developed within social psychology, it has major implications for the study of consumer behavior, and this is the primary focus of this paper. First, the basic postulates of self-perception theory are reviewed. Next, basic evidence for the theory is presented. Several current issues of relevance to consumer researchers are identified and research related to them are discussed. Finally, the implications of this work for consumer behavior theory and research are considered.

Self-Perception Theory

Self-perception theory (Bem, 1965, 1972) can be viewed as one portion of a general set of propositions commonly referred to as attribution theory. Like other attributional paradigms (see Kelley, 1967 on object attribution, Jones & Davis, 1965 on other person attribution, and Kelley, 1973 for an integrated view of all three paradigms), self-perception theory is phenomenological in character. That is, self-perception theory takes the perspective of the actor. It attempts to explain

and predict how individuals come to understand the causes of their own behavior, and to specify the consequences of this causal assignment.

The basic postulate of self-perception theory is that:

Individuals come to "know" their own attitudes, emotions, and other internal states partially by inferring them from observations of their own overt behavior and/or the circumstances in which this behavior occurs (Bem, 1972, p. 2).¹

Self-perception theory postulates the mechanism and the conditions under which a person uses his own behavior as data to make inferences about himself. For behavior to be used to infer individual characteristics, it must be a credible indicator of his internal state. In order to be credible, it must be perceived by the person as resulting from his internal motivations or true reactions to a stimulus, i.e., it must be attributed to properties of oneself as opposed to controlling aspects of the environment.² Thus, perceived causality of behavior is a primary determinant of the belief inference process.

Perception of causality, as indicated by Bem's proposition, depends partially upon the circumstances in which behavior is enacted. According to the discounting principle (Kelley, 1971), "the role of a given cause in producing a given effect is discounted if other plausible

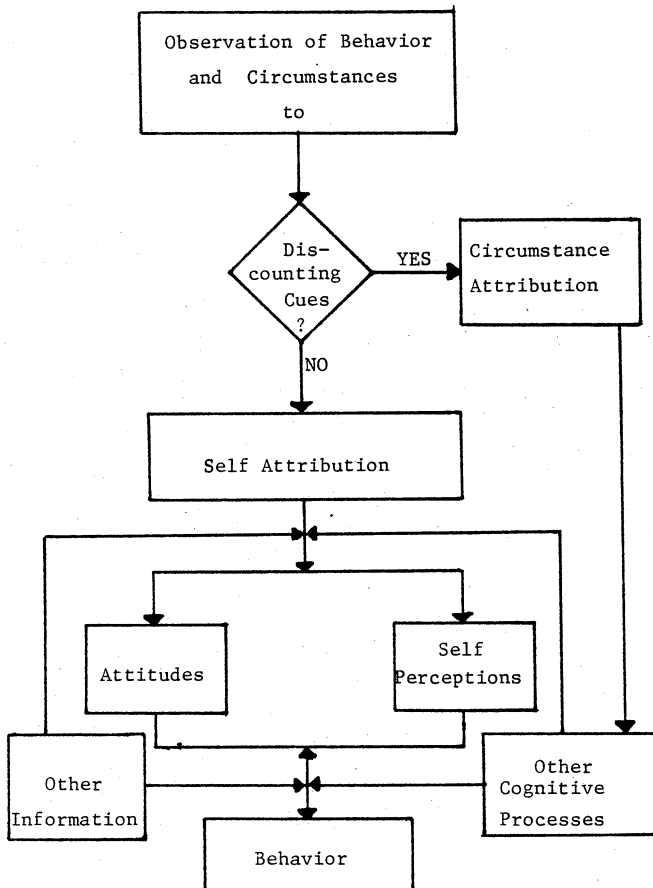
¹ A second postulate of self-perception theory is that individuals are functionally in the same position as an outside observer who must rely on external cues to infer the individual's inner state (Bem, 1972). While this proposition has caused a substantial amount of controversy because it implies that individuals have no more information about their feelings than observers, it has been shown to be tenuous at best (see Jones and Nisbett, 1971). Negation of this second postulate, of course, does not invalidate the first one which has received widespread support.

² A note on terminology should be made here. Many authors refer to this process as internal (self) versus external (situational contingency) attribution, while others (notably Nisbett and Valins, 1971) prefer the terms stimulus (perceived intrinsic properties) versus circumstance (factors external to the stimulus). The difference is more apparent than real. Behavior that is perceived to be elicited by the person's intrinsic feelings about the stimulus (stimulus attribution) implies internal motivation or causality. Behavior elicited by a desire to gain a particular reinforcement or to avoid punishment or by some other aspect of the circumstances (circumstance attribution) implies external causality. For simplicity, the terms "internal" and "external" will be used here.

ble causes are present" (p. 8).³ The presence of plausible external causal factors such as reward/cost contingencies, lack of volition, and the like should induce discounting of internal motivations as the cause of behavior. If a person performs an act for a large reward, for example, he would attribute his behavior to the reward rather than to his own positive attitude toward the act. When attributions are made to such external situational contingencies ("I was paid to do this"), no belief inferences would be made, and no link between behavior in one situation and behavior in another is expected. On the other hand, attributions to oneself ("I did this because I wanted to") are postulated to result in belief inferences ("I must believe in this issue") and subsequent behavior consistent with the initial behavior and internal attributions.

This attribution process is diagrammed in Figure 1. The relationship between past behavior, discounting cues, and future attitudes and actions has received the most empirical attention, and this body of literature provides the strongest support for the theory. The other parts of the process have received less attention, and these relationships are the source of current issues in self-perception research. The next section provides brief overview of this work.

FIGURE 1
A Schematic of the Attribution Process



³ Circumstantial factors can enhance the probability of self-attribution. Behavior enacted under inhibitory conditions (e.g., purchase at a premium price) should enhance the perception of internal motivation and result in strong belief inferences (augmenting principle, Kelley, 1971). These factors have been investigated only rarely, and thus are not discussed here.

Evidence relevant to self-perception theory can be discussed by first examining those studies which focus on the effects of past behavior, and then examining those that deal with the effects of perceived causality of behavior derived from Bem's (1972) central proposition. Several questions regarding the self-perception process emerging from these basic works and preliminary findings pertinent to them can then be discussed.

Effects of Past Behavior

At a minimum, two conditions must be demonstrated if the basic postulate of self-perception theory is valid. First, it must be shown that individuals do use their own behavior to infer their attitudes and/or to guide subsequent behavior. Second, it must be shown that circumstances surrounding the enactment of behavior affect this inference process.

The first of these conditions has been illustrated by several laboratory and field experiments in which a behavior is induced and subsequent attitudes and/or behavior are measured (e.g., Bandler, Madaras, & Bem, 1968; Freedman & Fraser, 1966; Pliner, Hart, Kohl, & Saari, 1974; Snyder & Cunningham, 1975; Valins, 1966). In a laboratory study (Valins, 1966), for example, male subjects were allowed to hear what was ostensibly their own heart beat while viewing pictures of nude females. Pictures associated with a (falsely manipulated) quickened heart rate were subsequently rated as more attractive and were frequently chosen as ones the subjects would like to keep than others. Similarly, in a field study utilizing what has become known as the foot-in-the-door paradigm, people were contacted at home and asked to display a small sign promoting a social cause in their window (Freedman & Fraser, 1966). Two weeks later, these individuals and a control group of subjects not contacted previously were asked to comply with another very large request. When both the issue and the task were similar across requests, 76% of the experimental participants as opposed to 16.7% of the control participants agreed to comply with the large request. Apparently, the initial behavior was self-attributed, causing subjects to draw belief inferences which enhanced the likelihood of subsequent compliance.

Effects of Perceived Causality

The second necessary pattern of data has been reported in numerous experiments which show that behavior enacted under discounting cue conditions does not persist in future time periods and may even decline in frequency (Bem, 1966; Davison & Valins, 1969; Calder & Staw, 1975; Deci, 1971; Kruglanski, Freedman, & Zeevi, 1971; Kruglanski, Alon, & Lewis, 1972; Kruglanski, Riter, Arati, Agassi, Montequio, Peri, & Peretz, 1975; Lepper, Greene, & Nisbett, 1973; Lepper & Green, 1975; Storms & Nisbett, 1970). Using an overjustification paradigm,⁴ for example, Lepper et al. (1973) found that children who expected and received a reward for engaging in a playtime activity subsequently spent less time engaging in that activity than a control group of children, even though they had previously evidenced considerable interest in it.

More recently, both of the two necessary conditions have been demonstrated in the consumer behavior literature. Scott (1976) asked individuals to take a two-week trial

⁴ Overjustification situations are the converse of insufficient justification ones, are those in which there are both internal and external motivations for engaging in an activity.

subscription to a weekly community newspaper either at the regular price (discounting cue absent) or at various discounted prices (discounting cues present). After the trial period had expired, these participants and a control group of subjects not receiving the trial offer were called and asked to subscribe on a regular basis. The results supported self-perception predictions in that large incentives for trial produced no more regular subscriptions than did the no-trial control condition. No or very small incentives, however, did have a positive effect on subscription behavior.

The negative effect of monetary incentives on subsequent behavior has been replicated in a social marketing context (Scott, in press) and in an interview request situation (Reingen & Kernan, 1977). And, Tybout (in press) and Dholakia & Sternthal (1977), using essentially the same foot-in-the-door paradigm, have shown that high credibility of the source of a message or request also acts as a discounting cue which attenuates subsequent behavior. Finally, Hansen (in review) has found that while monetary inducements to respond to a mail survey do not adversely affect response rate, they do result in lower quality of responses.

These studies provide considerable support for the self-perception thesis, and demonstrate the applicability of the self-perception framework for many consumer behavior contexts. However, they and other experiments conducted more recently raise some fundamental questions regarding the self-perception process and indicate areas where theoretical extensions and empirical research are needed to clarify our understanding of the link between past and future behavior. Six of these issues and preliminary data relevant to them will be discussed next.

Research Issues

Many areas of the self-perception process remain relatively unexplored. Following the diagram in Figure 1, these areas may be categorized as relating to the strength of the behavior-behavior link, the definition of past behavior, the relationship between the magnitude of discounting cues and behavior, the mediating role of attributions, the outcomes of self-perception processes, and the way in which experiential information is combined with other types of information. Research issues in each area and their implications will be briefly considered.

Issue 1: Strength of Self-Perception Effects

A very basic issue relates to the strength of the link between past and future actions. In their seminal study, Freedman and Fraser (1966) suggest that an initial behavior has a rather dramatic and persistent (two weeks) effect on subsequent actions. However, only behavioral intentions to comply with the large request were measured. Actual behavior was measured by Pliner et al. (1974), but the requests were made only one day apart, and difficulty of the behavior (e.g., donation to charity) was a dependent variable rather than an independent one. Scott (in press) varied second request task difficulty and measured both behavioral intentions and behavior in order to assess directly the strength of the initial behavior effect. The findings replicated those of Freedman and Fraser (1966) in that behavioral intentions to comply with the very difficult task were positively affected, but behavior was affected only when the second task was moderately difficult.

These studies show that an initial behavior does have a positive influence on subsequent behavior, but that this effect may be limited by the difficulty of that subsequent action. It seems likely that a series of requests could be used to gain compliance with very large requests, but this speculation has not been tested

empirically. Thus, the strength of influence strategies based on this paradigm is still unclear at least in the case where the two requests are made on two different occasions and where behavioral intentions are not synonymous with behavior. Other studies have examined the effects of size of the initial behavior on subsequent behavior, and these are discussed next.

Issue 2: What is Past Behavior?

In the field studies discussed previously (e.g., Freedman & Fraser, 1966; Pliner et al., 1974; Scott, 1976, in press), the initial behavior consisted of behavioral intentions to comply with a request. Two other types of experiments have varied this definition of past behavior. The first examines the effect of merely relabeling or reinterpreting some behavior that has already occurred, while the other focuses on enhancing the behavior-behavior link by increasing the level of difficulty of an initial, induced behavior.

Labeling studies manipulations are often no more than the provision of a belief-relevant label or interpretation for behavior that has already occurred. In a study by Tybout and Yalch (in review), for example, individuals responded to a survey about voting behavior. Half the individuals were then told that they were more likely to vote than the average person, while others were told that they were about average in their likelihood of voting. Examining records of a local election one week later, the authors found that more individuals given the "above average" feedback had indeed voted than individuals given the "average" feedback. These findings replicate those of Kraut (1973) who found that labeling individuals as charitable or uncharitable resulted in subsequent behavior consistent with these labels, and of Miller, Brickman, and Bolen (1975) who showed that labeling children as non-litterers or math achievers enhanced subsequent behavior relevant to that label. All of these results are consistent with self-perception assumptions about the cognitive status of attitudes and empirical evidence which indicates that people often do not know what their attitudes are (e.g., Ross, Insko, & Ross, 1971). They are also congenial with findings that attitudes are inferred from behavior only when behavior is perceived as belief-relevant, or belief expressive (Kiesler, Nisbett, & Zanna, 1969). That is, labeling affects behavior because it implies a link between past actions and beliefs, and makes this link salient to the individual.

All of these studies, however, have utilized a personal contact to deliver the label. Extending this work, Allen (1977) tested the relative effectiveness of a labeling message versus a typical persuasive message using television advertisements. The results were equivocal, but suggest that under some conditions people are sensitive to the explicit positive label ("You are the kind of person who...") of the labeling message and the implicit negative label of the persuasive message ("You ought to be..."). Too little is known about the qualifying conditions of labeling effects and the cognitive meaning of labels to reach definite conclusions, but the question of the situations in which labeling or persuasive messages are most effective is an intriguing one for consumer researchers.

A second type of study takes an approach opposite that of labeling, and seeks to test methods of strengthening the effect of an initial behavior by increasing its difficulty or size. Seligman, Bush, and Kirsch (1976), for example, found that very small behaviors had no effect on subsequent actions, while Pliner et al. (1974) found no differences between groups performing different levels of initial behaviors. It is clear from even a cursory examination of these discrepant findings that we actually know very little about what kind of an initial behavior is required for the self-perception process to become

operative, nor do we know along what dimensions this initial behavior should be classified. It may be that size of the request is not as important as degree of involvement, for example. These questions must be answered before influence strategies predicated on the paradigm can be used confidently.

Issue 3: Magnitude of Discounting Cues and Behavior

The definition of discounting cues as well as the definition of past behavior is an issue of particular relevance for consumer research. In many previous studies, a discounting cue is either present or absent. In contrast, consumer choice contexts often involve an initial behavior or trial of a product induced by a monetary incentive (e.g., coupons, cents-off deals, etc.) which may be of a variety of sizes. While self-perception theory does not specifically hypothesize a relationship between magnitude of the discounting cue and the likelihood of subsequent behavior, a reasonable derivation is that this relationship should be negative. Scott (1976; in press) addressed this issue empirically by including several levels of monetary incentives as treatments. Two findings are of interest. First, small incentives did not produce discounting in a commercial setting (Scott, 1976). Second, the relationship between magnitude of the incentive and subsequent behavior appeared to be negative, but not linear, in both a commercial (Scott, 1976) and a social issue setting (Scott, in press).

The first of these findings suggests that monetary incentives as discounting cues may have different effects in commercial than social settings. In commercial contexts, small incentives for trial may be perceived as so intrinsic to behavior that they are not perceived as salient causes of behavior. Kruglanski (1975) has shown that rewards intrinsic to the task (e.g., stock market choices) do not result in decreased interest or behavior as they do when the rewards are not intrinsically related to the task (e.g., a social game). Thus it may be that small monetary incentives do not lead to discounting of internal motivations in commercial contexts. It is only when they become very large, or when they are used in an unusual setting (e.g., a social program), that they are perceived as salient causes of behavior and discounting occurs. The findings of a single experiment should, of course, be regarded as the only tentative evidence at best. However, researchers need to be sensitive to areas where cross-situational generalizability may be problematic.

The second finding indicates that incentives as discounting cues may carry different meanings and thus activate different processes than other forms of discounting cues. That is, they not only provide contextual cues about behavior, but also specify the terms of an exchange in which perceptions of equity (Adams 1965) are relevant. When incentives are very large relative to the task, the person may feel that he has been inequitably rewarded and is obligated to do something in return to restore equity. An attempt to investigate this issue would be characterized by an experiment including several levels of incentives and measures of attitudes and perceptions of inequity as well as behavior. If the speculations about equity are correct, behavior would be curvilinearly related to levels of incentive, but attitudes would be negatively related and perceptions of inequity positively related to the magnitude of incentives. This type of research on the meaning and effects of monetary incentives is of central importance in understanding much marketplace behavior.

Issue 4: The Mediating Role of Attributions

Many self-perception studies merely assume the presence of an attributional mechanism from behavioral findings

consistent with attribution predictions. However, varying the presumed antecedents of self-perceptions and observing the outcomes is not sufficient to demonstrate the efficacy of the theory (see Burnkrant, 1975). Direct assessment of attribution is necessary since competing theories often make the same behavioral predictions.

Among studies which do include measures of cognitive processes, a frequent finding is significant behavioral effects, but no or only slight effects on cognitive variables assumed to mediate behavior (e.g., self-perceptions, attitudes, attributions). An interesting theoretical rationale for these results is one proposed by Nisbett and Valins (1971) who postulate that self-attributions and the corresponding self-perceptions or beliefs have a cognitive status similar to an hypothesis. That is, performance of a behavior causes individuals to attribute it to themselves, and this attribution in turn causes the individual to form an hypothesis about his beliefs, which is then tested out in subsequent actions. The behavioral effects demonstrated in previous studies then result from a desire to test this attributional hypothesis by gaining more information rather than to act consistent with stable attributions and attitudes.

The hypothesis-testing proposition is compatible with the results of several studies. For example, Barefoot and Straub (1971) were able to replicate Valin's (1966) false heartrate feedback findings only when subjects were given sufficient time to examine the pictures. Apparently the quickened heartrate caused subjects to view the pictures more closely and confirm that they were indeed more attractive. Similarly, Valins (1974) and Ross, Lepper, and Hubbard (1975) have demonstrated that merely debriefing subjects about the falsity of the feedback had no effect. Again, it seems that feedback stimulates an individual to gather more information to confirm the belief hypothesis. Once confirmed, the falsity of the feedback is irrelevant to the belief.

A recent experiment tested this attribution process directly (Scott & Yalch, in press). Subjects were asked to test a new soft drink under incentive or no incentive conditions. After making this choice, half the subjects were allowed to visually examine the product. All subjects then tasted the product which was manipulated to provide either a good, neutral, or poor experience. When subjects were allowed to examine the product, the familiar discounting results were found in the neutral and poor taste conditions; incentives undermined subjects' attitudes toward the drink, behavioral intentions to purchase it, and actual choice of it relative to no-incentive conditions. The reverse effect of incentive was found when no examination opportunity was given. It appears that individuals do infer their beliefs from their own behavior, but only when they have an opportunity to persuade themselves of the veracity of the experimental information.

The hypothesis-testing proposition and related empirical work represent beginning attempts to outline the conditions under which attributional hypotheses are translated into beliefs and behavior. Unfortunately, this begs the question of what starts the attribution process in the first place. That individuals do not always engage in causal analyses is suggested by findings from Tybout (in press). In her study, the predicted discounting cue effects were obtained only when subjects' attention was drawn to their behavior by a question asking why they behaved as they did.

The issue is not really whether individuals do or do not engage in attributional analyses, but rather is "under what conditions do they perform these analyses." The answer may be found in terms of individual difference characteristics (e.g., absence of self-schema on a dimension as in Markus (1977), internal versus external

locus of control as in Shaver (1976) or low education as in Tybout, in press). Or, it may be found in terms of the characteristics of the context of behavior (e.g., expectation of future consequences as suggested by Berscheid, Graziana, Monson, & Denmer, 1976, and relevance of the behavior to beliefs as in Kiesler, Nisbett, & Zanna, 1969). Unfortunately, little evidence exists to answer this question conclusively although the literature certainly provides some hypotheses. It is critical that future research address this issue since self-perception theory will have little applicability to consumer behavior if it turns out that only a small segment of the population engage in these analyses spontaneously and that there are no methods to encourage such analysis. In sum, the external validity of the self-perceptions process must be defined and demonstrated.

Issue 5: The Cognitive Consequences of Self-Attributions

Related to the issue of when attributions are made and their cognitive status once they are made is the issue of what happens as a result of self-attributions. Much evidence exists which demonstrates a behavior change, but we know very little about the cognitive changes that occur. In interpreting their results, Freedman and Fraser (1966) offered the following possibilities:

Once he has agreed to a request, his attitude may change. He may become in his own eyes, the kind of person who does this sort of thing, who agrees to requests made by strangers, and who takes action on things he believes in, who cooperates with good causes (p. 201).

Thus behavior may change a person's perception of himself in general, or may alter attitudes specifically related to the behavior, or both. And, there is some evidence to suggest that other cognitive activities occur as well.

With regard to the self-perception versus specific attitude question, a few studies indicate that one's self-perception is most affected. Freedman and Fraser (1966), for example, reported positive (but not significant) effects of past behavior even when the issue and task were different across requests. Similarly, significant positive effects of behavior on subsequent actions have been observed when the issue was different across requests (Snyder & Cunningham, 1975). These studies provide indirect evidence that issue-specific attitude change is not solely responsible for the behavioral effects.

Direct evidence for an effect on self-perception is found in Scott (in press). Attitudes and perceived personal activism were both measured in this study, with larger (although not significant) effects found on activism than on issue-specific attitudes. Some methodological problems exist in this design, however, and the results must be viewed as tentative. And, attitudes of the subject population were already favorable, so a ceiling effect may have been encountered.

Future research is needed which clarifies the cognitive effects of self-attributions since this will be a critical factor in determining the applicability of self-perception based strategies to consumer settings. If issue-specific attitudes are not affected, the applicability of the theory to choice decisions among similar alternatives is problematic. On the other hand, there are also many situations in which attitudes are already favorable but behavioral enactment of these attitudes is low (e.g., everyone believes that conservation of energy is important, but few take action), and changes in self-perceptions may provide an energizing force on behavior.

Finally, there is evidence to suggest that the attribution process may affect other cognitive variables indirectly. Several studies show that providing feedback about one's behavior, or making that behavior salient instigates a search for information (e.g., Taylor, 1975; Berscheid et al., 1976). And, other experiments indicate that the provision of rewards or incentives for behavior also have this effect. Using a Bayesian framework for analyzing the value of information, Yalch and Scott (in review) and Scott and Tybout (1977) have found that subjects who agreed to test a product under incentive (reward) conditions responded more to information provided them after this choice.

Issue 6: Combining Experiential and Other Types of Information

Self-perception theory offers fairly clear specifications of when an individual uses his own behavior to formulate attitudes and guide behavior. Other attributional paradigms (e.g., object and other person) suggest other types of information that may be used to form judgments. Specifically, distinctiveness of response, consistency of response, and consensus information may be used when the person forms judgments over time (see Kelley, 1976; 1973). It seems likely in most consumer contexts that individuals will have access to at least some of these types of information as well as to information from their own experiences. We know that self-perception induced beliefs can affect both the quantity of information and the perception of the content of that information at least when it is somewhat ambiguous (e.g., tasting a "new" soft drink as in Scott & Yalch, in press). The question now is whether these beliefs will be given greater or lesser weight than information of another type.

Burnkrant and Cousineau (1975) have shown that individuals' judgments of product quality are influenced by the degree of consensus exhibited by others. The greater the apparent consensus, the more an individual is influenced by this information. These authors did not, however, examine the relative influence of experiential and consensus information. That is, they did not induce a self-perception based belief and observe changes in those beliefs as a function of consensus information. Research by Nisbett, Borgida, Crandell, and Reed (1976) indicates that individuals often ignore consensus information and rely more heavily on their own experiential data, and Scott and Tybout (1977) found that the timing of consensus information presentation as well as whether discounting cues are present affect the acceptance of consensus information. That is, rewarded subjects showed a primacy effect in that they were more influenced by consensus information when it preceded actual product experience. Unrewarded subjects generally were less responsive to both consensus and actual product experience.

The process by which information sources are combined is undoubtedly a complex one which will require much research to unravel. It seems clear, however, that a merging of the paradigms is necessary to extend the external validity of the research questions posed and the experimental designs used in consumer behavior research.

Implications and Conclusions

Self-perception theory and research promises to add much to our understanding of how consumers learn from their own experiences and the consequences of this learning for future actions. The phenomenological focus of this framework forces one to examine not only behavior itself, but also to examine how the individual perceives the behavior. The meaning an individual assigns to his experiences is of critical importance in predicting and explaining subsequent actions. Self-perception theory

acknowledges that learning from past behavior is not a simple, automatic response. Rather, it is a complex process in which what is learned from previous experiences is dependent upon many interacting factors. While situational factors affect the perception of the outcomes of consumer behavior, for example, it is also likely that some outcomes will affect the perception of situational cues.

Self-perception theory (and attribution theory in general) is inherently dynamic, and thus it orients researchers to examine behavior over time. Situational characteristics, in self-perception terms, affect behavior in both immediate and future time periods. Investigation of only the immediate impact of situational variables tells only half the story. Similarly, strategies designed to modify behavior must be evaluated at several points in time since their long-run effects may vary from their short term ones.

Caution must be exercised, however, in expecting too much of self-perception theory as developed thus far. Consumer researchers must be aware of the unresolved issues surrounding the self-perception process, and must be vigilant for conditions which may limit its cross-situational generalizability. As noted in this paper, we have only just begun to examine the self-perception process in any depth. And, consumer researchers must be careful even in applying the knowledge gained in other contexts. In consumer contexts, it may be that some product decisions involve attributional processes more frequently than others, that some individuals utilize attributional analyses more than others, or that special strategies must be developed to encourage attributional processes.

It should be clear from this paper that much research remains to be done in explicating the self-perception process and its implications for consumer behavior. Given its potential for shedding light on that ubiquitous feedback loop, however, such research seems well worth our time and resources.

References

- J. Stacy Adams, "Inequity in Social Exchange," in L. Berkowitz (ed.), Advances in Experimental Social Psychology, Vol. 2 (New York: Academic Press, 1965), 267-299.
- Chris T. Allen, "Delimiting and Describing the Phenomenon of Socially-Conscious Consumption Behavior: An Experimental Approach," Unpublished dissertation, The Ohio State University, 1977.
- R. G. Bandler, G. R. Madaras, and D. J. Bem, "Self-Observation as a Source of Pain Perception," Journal of Personality and Social Psychology, 9 (July 1968), 205-209.
- J. C. Barefoot and R. B. Straub, "Opportunity for Information Search and the Effect of False Heart-Rate Feedback," Journal of Personality and Social Psychology, 17 (February 1971), 154-157.
- Frank M. Bass, "The Theory of Stochastic Preference and Brand Switching," Journal of Marketing Research, 11 (February, 1974), 1-20.
- D. J. Bem, "An Experimental Analysis of Self-Persuasion," Journal of Experimental Social Psychology, 1 (August, 1965), 199-218.
- D. J. Bem, "Self-Perception Theory," in L. Berkowitz (ed.), Advances in Experimental Social Psychology, Vol. 6 (New York: Academic Press, 1972), 1-62.
- D. J. Bem, "Inducing Belief in False Confessions," Journal of Personality and Social Psychology, 3 (June, 1966), 707-710.
- Ellen Berscheid, William Graziano, Thomas Monson, and Marshall Dermer, "Outcome Dependency: Attention, Attribution, and Attraction," Journal of Personality and Social Psychology, 34 (November, 1976), 978-989.
- Robert E. Burnkrant, "Attribution Theory in Marketing Research: Problems and Prospects," in M. J. Schlinger (ed.), Advances in Consumer Research, Vol. 2 (Association for Consumer Research, 1975), 465-469.
- Robert E. Burnkrant and Alain Cousineau, "Informational and Normative Social Influence in Buyer Behavior," Journal of Consumer Research, 2 (December, 1975), 206-215.
- B. J. Calder and B. M. Staw, "Self-Perception of Intrinsic and Extrinsic Motivation," Journal of Personality and Social Psychology, 31 (April, 1975), 599-605.
- G. Davison and S. Valins, "Maintenance of Self-Attributed and Drug-Attributed Behavior Change," Journal of Personality and Social Psychology, 11 (January, 1969), 25-33.
- E. L. Deci, "The Effects of Externally Mediated Rewards on Intrinsic Motivation," Journal of Personality and Social Psychology, 18 (April, 1971), 105-115.
- Ruby Roy Dholakia and Brian Sternthal, "Highly Credible Sources: Persuasive Facilitators or Persuasive Liabilities?" Journal of Consumer Research, 3 (March, 1977), 223-232.
- James F. Engel, David T. Kollat, and Roger D. Blackwell, Consumer Behavior, (New York: Holt, Rinehart, and Winston, 1973).
- J. L. Freedman and S. C. Fraser, "Compliance Without Pressure: The Foot-in-the-Door Technique," Journal of Personality and Social Psychology, 4 (August, 1966), 195-202.
- Robert A. Hansen, "The Effects of Monetary and Non-Monetary Incentives on Mail Survey Respondent Behavior: A Self-Perception Interpretation," in review.
- John A. Howard and Jagdish N. Sheth, The Theory of Buyer Behavior (New York: John Wiley, 1969).
- E. E. Jones and K. E. Davis, "From Acts to Dispositions," in L. Berkowitz (ed.), Advances in Experimental Social Psychology, Vol. 2 (New York: Academic Press, 1965), 219-266.
- H. H. Kelley, "Attribution Theory in Social Psychology," in D. Levine (ed.), Nebraska Symposium on Motivation (Lincoln, Nebraska: University of Nebraska Press, 1967).
- H. H. Kelley, "Attribution in Social Interaction," in E. E. Jones et al., Attribution: Perceiving the Causes of Behavior (Morristown, New Jersey: General Learning Press, 1971), 151-173.
- H. H. Kelley, "The Processes of Causal Attribution," American Psychologist, 38 (February 1973), 107-128.
- C. A. Kiesler, R. E. Nisbett, and M. Zanna, "On Inferring One's Beliefs From One's Behavior," Journal of Personality and Social Psychology, 11 (April, 1969), 321-327.
- R. E. Kraut, "Effects of Social Labeling on Giving to Charity," Journal of Experimental Social Psychology, 9 (November, 1973), 551-562.

- A. W. Kruglanski, "The Endogenous-Exogenous Partition in Attribution Theory," Psychological Review, 82 (November, 1975), 387-406.
- A. W. Kruglanski, S. Alon, T. Lewis, "Retrospective Misattribution and Task Enjoyment," Journal of Experimental Social Psychology, 8 (November, 1972) 493-501.
- A. W. Kruglanski, I. Freedman, and G. Zeevi, "The Effect of Extrinsic Incentive on Some Qualitative Aspects of Task Performance," Journal of Personality, 39 (December, 1971), 606-617.
- A. W. Kruglanski, A. Riter, D. Arazi, R. Agassi, J. Monteqio, I. Peri, and M. Peretz, "Effects of Task Intrinsic Rewards Upon Extrinsic and Intrinsic Motivation," Journal of Personality and Social Psychology, 3 (April, 1975), 699-705.
- A. A. Kuehn, "Consumer Brand Choice as a Learning Process," Journal of Advertising Research, 2 (December, 1962), 10-17.
- M. R. Lepper and D. Greene, "Turning Play into Work: Effects of Adult Surveillance and Extrinsic Rewards on Children's Intrinsic Motivation," Journal of Personality and Social Psychology, 31 (March, 1975), 479-486.
- M. R. Lepper, D. Greene, and R. E. Nisbett, "Undermining Children's Intrinsic Interest With Extrinsic Reward: A Test of the Overjustification Hypothesis," Journal of Personality and Social Psychology, 28 (October, 1973), 129-137.
- Hazel Markus, "Self Schemata and Processing Information About the Self," Journal of Personality and Social Psychology, 35 (February, 1977), 63-78.
- R. L. Miller, P. Brickman, and D. Bolen, "Attribution Versus Persuasion as a Means of Modifying Behavior," Journal of Personality and Social Psychology, 31 (March, 1975), 430-41.
- Francesco M. Nicosia, Consumer Decision Processes (Englewood Cliffs, New Jersey: Prentice-Hall, 1966).
- R. E. Nisbett, E. Borgida, R. Crandall, and H. Reed, "Popular Induction: Information is not Always Informative," in J. Carroll and J. Payne (eds.), Cognition and Social Behavior (Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1976).
- R. E. Nisbett and S. Valins, "Perceiving the Causes of One's Own Behavior," in E. E. Jones et al., Attribution: Perceiving the Causes of Behavior (Morristown, New Jersey: General Learning Press, 1971), 63-76.
- P. Pliner, H. Hart, J. Kohl, and D. Saari, "Compliance Without Pressure: Some Further Data on the Foot-in-the-Door Technique," Journal of Experimental Social Psychology, 10 (January, 1974), 17-22.
- P. H. Reingen and J. B. Kernan, "Compliance With an Interview Request: A Foot-in-the-Door, Self-Perception Interpretation," Journal of Marketing Research, 14 (August, 1977), 365-369.
- M. Ross, C. A. Insko, and H. S. Ross, "Self-Attribution of Attitude," Journal of Personality and Social Psychology, 17 (March, 1971), 292-297.
- L. Ross, M. Lepper, and M. Hubbard, "Perserverance in Self-Perception and Social Perception: Biased Attributional Processes in the Debriefing Paradigm," Journal of Personality and Social Psychology, 32 (November, 1975), 880-892.
- Carol A. Scott, "The Effect of Trial and Incentives on Repeat Purchase Behavior," Journal of Marketing Research, 13 (August, 1976), 263-269.
- Carol A. Scott, "A Self-Perception Approach to Modifying Socially-Conscious Behavior," Journal of Consumer Research, in press.
- Carol A. Scott and Alice M. Tybout, "The Effects of Product Trial, Incentives, and Consensus Information on Consumer Choice," working paper, The Ohio State University, 1977.
- Carol A. Scott and Richard F. Yalch, "A Test of the Self-Perception Explanation of the Effects of Rewards on Intrinsic Interest," Journal of Experimental Social Psychology, in press.
- C. Seligman, M. Bush, and K. Kirsch, "Relationship Between Compliance in the Foot-in-the-Door Paradigm and Size of First Request," Journal of Personality and Social Psychology, 33 (May 1976), 517-520.
- Kelly Shaver, An Introduction to Attribution Processes (Cambridge, Massachusetts: Winthrop, 1975).
- Mark Snyder and Michael R. Cunningham, "To Comply or Not Comply: Testing the Self-Perception Explanation of the 'Foot-in-the-Door' Phenomenon," Journal of Personality and Social Psychology, (January, 1975), 64-67.
- Michael D. Storms and R. E. Nisbett, "Insomnia and the Attribution Process," Journal of Personality and Social Psychology, 16 (October, 1970), 319-328.
- Shelley E. Taylor, "On Inferring One's Attitudes from One's Own Behavior: Some Delimiting Conditions," Journal of Personality and Social Psychology, 31 (January, 1975), 126-131.
- Alice M. Tybout, "The Relative Effectiveness of Three Behavioral Influence Strategies as Supplements to Persuasion in a Marketing Context," Journal of Marketing Research, in press.
- Alice M. Tybout and Richard F. Yalch, "On Getting Out the Vote: Effects of Labeling on Voter Turnout," in review.
- S. Valins, "Cognitive Effects of False Heart-Rate Feedback," Journal of Personality and Social Psychology, 4 (October, 1966), 400-408.
- S. Valins, "Persistent Effects of Information About Internal Reactions: Ineffectiveness of Debriefing," in H. London and R. E. Nisbett (eds.), Thought and Feeling: Cognitive Modification of Feeling States (Chicago: Aldine, 1974).
- Richard F. Yalch and Carol A. Scott, "Consumer Response to Initial Product Trial: A Bayesian Analysis," in review.

A TWO-STAGE THEORY OF INFORMATION PROCESSING IN PERSUASION: AN
INTEGRATIVE VIEW OF COGNITIVE RESPONSE AND SELF-PERCEPTION THEORY

Alice M. Tybout, Northwestern University
Brian Sternthal, Northwestern University
Bobby J. Calder, Northwestern University

Abstract

Cognitive response theory is presently not a sufficient explanation of the information processing underlying persuasion. Initial opinion, one of the key constructs in cognitive response theory, is not well specified. This paper uses self-perception theory to specify the nature of the initial opinion construct. This integrative view emphasizes the importance of message related behaviors as cues utilized by consumers.

Introduction

During the last decade cognitive response theory has emerged as the most viable explanation of how individuals process persuasive information. The theory, however, is not a sufficient explanation of the persuasive process; it relies on the construct of initial opinion without fully specifying the nature of these opinions. Self-perception theory complements the cognitive response view, suggesting the process by which individuals acquire initial opinions. In this paper a common language is developed which links these theories. This integration provides a comprehensive theory which is sufficient to account for persuasion. In this integrative theory, initial opinion is a fully specified construct. An important implication of the theory is that persuasion depends on cues provided by message related behaviors as well as by the message itself.

This paper is divided into three sections. First, cognitive response theory extended to include memory structure is described. Next, an integrated theory is developed by incorporating the self-perception process into the cognitive response view. In the final section, implications of the comprehensive theory are outlined.

Extended Cognitive Response Theory

According to the cognitive response formulation, much of persuasion is self-persuasion. In response to a persuasive communication, individuals generate and rehearse their own repertoire of thoughts as well as those included in the information presented. The thoughts generated and rehearsed depend to a large extent on information recipients' initial opinion toward the position being advocated (cf. Greenwald, 1968). Those who are initially opposed to an advocacy are likely to rehearse counterarguments to it. And the more extensively people rehearse counterarguments, the more likely these thoughts are to become part of their attitudinal disposition toward the issue, and the less likely people are to be persuaded to adopt the advocated position. In contrast, those individuals who have a favorable initial opinion toward an issue are likely to generate and rehearse support arguments, thus enhancing the probability of their being persuaded by the information presented. (Petty, 1977; Sternthal, Dholakia and Leavitt, in review).

The description of cognitive response as a mediator of persuasion underscores the theoretical importance of the initial opinion construct. Initial opinion guides the generation of thoughts that ultimately affect persuasion.

However, for initial opinion to have integrity as a construct, the mechanism by which it influences processing must be described rather than merely stated. This can be achieved by extending cognitive response to include concepts related to memory structure (cf. Calder, 1978). More specifically, this view suggests that processing entails a short-term memory store and a long-term memory store. Short-term memory holds information which is being processed actively. Long-term memory holds most, if not all, of the information an individual has ever processed. To proceed further, the information held in long term memory must be retrieved and transferred to short term memory. In terms of cognitive response, memory structure implies that information from a message as well as other incoming information is initially represented in short-term memory as cognitive responses. In turn, these cognitive responses trigger the retrieval of further information from long-term memory and its registration in short-term memory.

An Integrative Theory

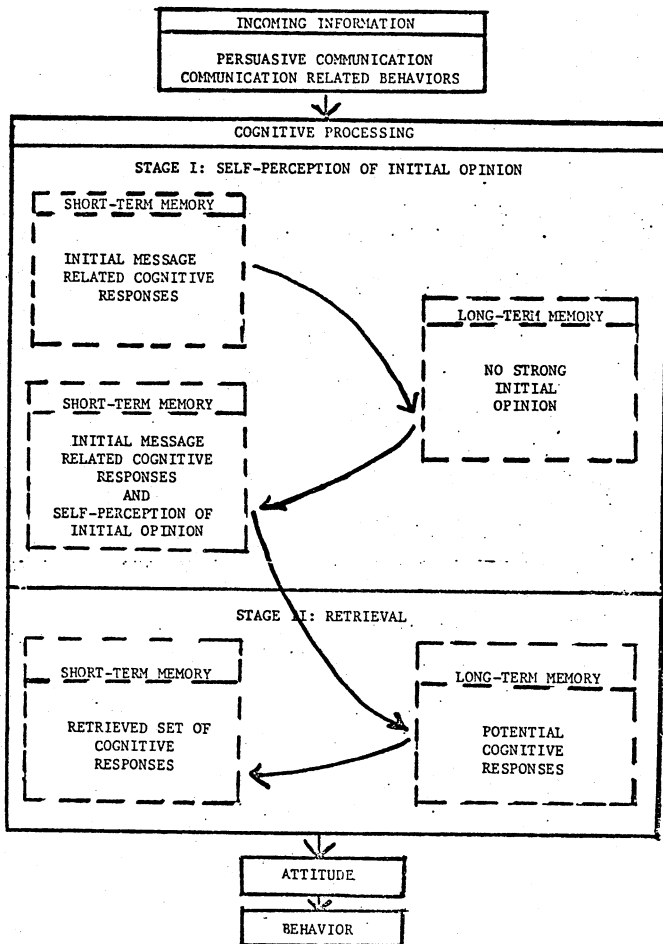
The extended cognitive response theory presented above implies that the role of initial opinion in current cognitive response theory may be equated with retrieval from long-term memory. The initial cognitive responses in short-term memory determine whether long-term memory will be searched for support or counterargument thoughts. It may be appropriate to think of persuasion as a two-stage process. In the first stage, incoming information is represented more or less faithfully in short-term memory. If a person already has strong general thoughts in long-term memory (perhaps as the product of self-perception), these thoughts will be triggered at this point and also entered into short-term memory. It is these cognitive responses which form an initial opinion that guides further retrieval from long-term memory. This subsequent retrieval is the second stage of the process.

Consider a second and more problematic case for cognitive response theory. What occurs if no strong general thoughts are already stored in long-term memory? That is, what if the message triggers no existing initial opinion. Cognitive response theory is not sufficient to explain the persuasive process in this instance, unless it is bolstered by self-perception theory. This latter theory describes the process by which individuals infer the causes of their own behavior when they do not have strong initial opinions. These inferences then guide subsequent action. When a person cannot attribute his/her actions to external causes, an internal disposition is inferred (e.g. liking a product). This disposition, in turn, may guide subsequent action. When a person's behavior is attributable to external causes such as social pressure or monetary incentives, no inference about internal states will be drawn with certainty. As a result, a disposition toward the behavior is unlikely to be formed and behavior will persist only in the presence of the external cause. Support for these self-perception predictions has emerged both in studies using the foot-in-the-door technique (cf. Freedman and Fraser, 1966;

Pliner, Hart, Kohl, and Saabi, 1974; Reingen and Kernan, 1977; Snyder and Cunningham, 1975; Scott, 1976; Tybout, 1978) and labeling (cf. Kraut, 1973; Miller, Brickman, and Bolen, 1973; Tybout and Yalch, in review).

In situations where the message triggers no initial opinion, individuals will examine behaviors associated with the message to infer their initial opinion (Figure 1, Stage I). The self-perception process ensues; if no external cause for the behavior is salient, an initial opinion consistent with the behavior is inferred. This opinion guides Stage II of the process (Figure 1). A final set of cognitive responses is retrieved from all of the potential cognitive responses in long-term memory.

FIGURE 1
SCHEMATIC DIAGRAM OF THE INTEGRATED THEORY
For the Case of No Initial Opinion



In an experiment pertaining to the persuasive effect of source credibility, Dholakia and Sternthal (1977) presented subjects with a communication requesting their compliance with a request attributed to either a high or low credibility source. Subjects' compliance was measured either before or after the administration of an attitude battery. Because message recipients were unfamiliar with the issue, it was expected that the highly credible source would induce a more positive attitude toward his advocacy than a low credibility source when attitude measures were administered prior to behavior. In this situation, a high credibility source was more likely to inhibit the retrieval and transfer of counterarguments from long to short-term memory than a low credibility source. Of greater interest in the present context is the source credibility effect when subjects behavior (i.e. compliance with the request) was available as a cue. Here, it was found that the low credibility source was more persuasive than the highly credible communicator.

To explain this result, we must invoke the common language. Consider for example those subjects who complied with the communicator's request. When the source was highly credible, individuals were likely to consider this stimulus as a likely cause of their compliance and therefore discount personal reasons. The relatively neutral initial opinion resulting from this attributional work would thus stimulate the retrieval and transfer of few support arguments and counterarguments to short-term memory when subjects were called upon to express their attitude toward the issue advocated. As a result, subjects' attitude would not be highly favorable or highly unfavorable. In contrast, those who complied in the low credibility condition could only plausibly attribute that behavior to a positive initial opinion toward the issue. In turn, this positive initial opinion would stimulate the retrieval and transfer of arguments supporting the advocacy to short-term memory when subjects were subsequently asked to complete the attitude questionnaire. Thus, attitude would be highly positive. These predictions were confirmed by Dholakia and Sternthal (1977).

Conclusion

As noted at the outset of this paper, cognitive response theory is possibly not a sufficient explanation for persuasion. The driving force in cognitive response theory is initial opinion. Yet the source of this opinion is not fully specified. While consumers may have stored opinions about some products/issues (possibly derived on the basis of their past behavior, circumstances surrounding that behavior etc.), for many products/issues no strong stored opinion may exist. In such instances, the individuals' self-perception of any message-related behaviors may provide the driving force for information processing. Our integration of cognitive response and self-perception theory provides a model which accounts for information processing regardless of whether or not an initial opinion exists. (See Table 1).

This integrative view is perhaps most important in underscoring the need to go beyond examining just the message in consumer research. Equally important may be the behavior and the perceived causes and significance of behavior associated with the message. On this view, consumer research has been myopic in investigating the non-message cues utilized by consumers.

In sum, we are suggesting that by developing a common language to link cognitive response and self-perception we can explain consumer behavior phenomena that cannot be adequately explained by either theory alone. Self-perception does not address the mechanism by which individuals process information. This is described by cognitive response. On the other hand, cognitive response does not consider a key source of initial opinion which may mediate persuasion, namely message related behaviors. This is addressed by self-perception theory. The importance of developing a common language may be illustrated with a recent investigation.

TABLE 1
POSSIBLE STATUS OF INITIAL OPINION

CASE	HYPOTHESIZED PROCESSING
Pre-existing initial opinion	Retrieved from long-term memory and used in retrieval of subsequent cognitive responses
No initial opinion, self-perception possible	Self-perception based on message related behaviors used in retrieval of subsequent cognitive responses
No initial opinion, self-perception not possible (i.e. behavior discounted)	"Random" retrieval

Alice M. Tybout and Richard F. Yalch, "On Getting out the Vote: Effects of Labeling on Voter Turnout," in review.

Alice M. Tybout, "The Relative Effectiveness of Three Influence Strategies as Supplements to Persuasion in a Marketing Context," Journal of Marketing Research, (February, 1978).

References

- Bobby J. Calder, "Cognitive Response, Images, and Scripts: What is the Cognitive Basis of Attitude?" Advances in Consumer Research, Volume 5, 1978.
- Ruby R. Dholakia and Brian Sternthal, "Highly Credible Sources: Persuasive Facilitators or Persuasive Liabilities," Journal of Consumer Research, 3(March, 1977) 223-232.
- Johnathan L. Freedman and Scott C. Fraser, "Compliance Without Pressure: The Foot-in-the-Door Technique," Journal of Personality and Social Psychology, 4(August, 1966), 195-202.
- Anthony Greenwald, "Cognitive Learning, Cognitive Response to Persuasion, and Attitude Change," in Anthony Greenwald, Timothy Brock, and Thomas Ostrom (eds.), Psychological Foundations of Attitudes, New York: Academic Press, 1968, 147-170.
- Robert E. Kraut, "Effects of Social Labeling on Giving to Charity," Journal of Experimental Social Psychology, 9(1973) 551-562.
- R.L. Miller, P. Brickman and D. Boten, "Attribution versus Persuasion as a Means for Modifying Behavior," Journal of Personality and Social Psychology, 31(March, 1975) 430-441.
- Richard Petty, "The Importance of Cognitive Responses in Persuasion," Advances in Consumer Research, vol. 4, 1977, 357-362.
- Patricia Piner, Heather Hart, Joanne Kohl and Dory Saabi, "Compliance Without Pressure: Some Further Data on the Foot-in-the-Door Technique," Journal of Experimental Social Psychology, 2(November, 1965), 669-676.
- Peter Reingen and Jerome Kernana, "Compliance with an Interview Request: A Foot-in-the Door, Self-Perception Interpretation," Journal of Marketing Research 14(August, 1977) 365-369.
- Carol A. Scott: "The Effects of Trial and Incentives on Repeat Purchase Behavior," Journal of Marketing Research, 13(August, 1976), 263-269.
- Mark Snyder and Michael R. Cunningham, "To Comply or Not Comply: Testing the Self-Perception Explanation of the 'Foot-in-the-Door' Phenomenon," Journal of Personality and Social Psychology, 31(January, 1975), 64-67.
- Brian Sternthal, Ruby R. Dholakia, and Clark Leavitt, "The Persuasive Effect of Source Credibility: Tests of Cognitive Response," in review.

*
CUE UTILIZATION IN PRODUCT PERCEPTION

Robert E. Burnkrant, University of California, Berkeley

Abstract

Cue utilization in product perception is viewed as complex information processing. It involves a process of making inferences about products from the configuration of cues available. It is maintained that this type of information processing is more likely to be engaged in with complex rather than simple products. It is also suggested that the consistency in meaning of cues associated with a product is directly related to the intensity of the perception of that product. Preliminary results of a study to test these hypotheses are briefly reviewed. The relevance of attribution theory to product perception is discussed.

Much of the information people acquire about products comes from the processing of product related cues. These cues provide a basis for making inferences about the product and give it meaning which frequently goes beyond and departs from the specific arguments made in advertising appeals. As such, they become important determinants of how people behave toward products.

An individual in a purchasing context is frequently faced with the problem of evaluating the worth of a product which he cannot possibly objectively evaluate from physical manipulation and observation of its characteristics alone. Among the strategies employed by the consumer is processing information contained in messages (i.e., advertisements) about the product. This has been given considerable attention in the literature. Much less attention has been devoted to another strategy. The consumer may turn to cues given by the product, the context in which it is found and the people who use it. These cues can provide a basis for inferring the characteristics of the object under consideration. This information processing involves making inferences about the product from the selective processing of product related cues.

It is important for marketers to come to grips with some of the factors which will determine the extent to which inferences will be made in particular situations and to identify cues people utilize when they make inferences. This will permit marketers to employ cues more consciously and appropriately with products and in situations in which these cues are likely to have an affect on consumers' product evaluations. This requires that we determine which types of products and situations will lead to cue processing. It requires a concerted effort to determine which product and other related factors are used by people as cues. It is also necessary to consider how various combinations of these cues will affect the inferences people make. The study to be described here addresses two of these issues. It attempts to identify a product dimension which will be directly related to the use of cues in product evaluation, and it attempts to assess the affects of variations in cue consistency on product evaluation.

The author would like to acknowledge the support of the Center for Research in Management Science, the Committee on Research and the Institute of Business and Economic Research of the University of California, Berkeley. He would also like to thank Mr. Scott Vitell for research assistance provided this project.

The process of inferring the characteristics of an object from indirect information is clearly more complex (i.e., involves more cognitive effort) than the direct processing of observable product attributes. As a result, it would be expected to be engaged in only when available direct information is insufficient to permit satisfactory performance on the impending task. The assessment of product differences in cue utilization naturally turns, then, to a consideration of dimensions which will distinguish between products requiring the individual to infer from indirect information and those which do not require this type of information processing.

Product familiarity has been found to be related to cue utilization in product evaluation. Monroe (1976) found that price played a greater role when subjects were unfamiliar with the brand than when they were familiar with it. Similar findings were also obtained by Smith and Broome (1966). In Jacoby, Olson and Haddock (1971) the influence of price ceased when price was paired with an already known brand.

It is suggested that if we control for familiarity, the tendency to make inferences from such cues as price should be directly related to product complexity. What is meant here by product complexity is the typical or mean number of dimensions by which the product is represented in the minds of its potential customers. It is expected that some products are largely represented in only very few, perhaps one or two, dimensions. These products are considered to be relatively simple. On the other hand, some products are likely to be represented by most people in several dimensions, and these products are considered to be complex.

This treatment of product complexity is adapted from work focusing on cognitive complexity. Researchers concerned with this topic have treated cognitive complexity as an individual difference variable and have made predictions about the effects of complexity on information processing (e.g., Kelly, 1957; Schroder, Driver and Streufert, 1967). It has been found that "most persons are relatively 'complex' in some areas of cognition and relatively 'simple' in others" (Gardner and Schoen, 1962). It follows that most people may be "complex" with regard to some products and "simple" with regard to other products. If this is true this distinction should have implications for information processing, including the tendency to base a products' perception on inferences.

Schroder, Driver and Streufert (1967) point out that "the more judgmental space contains dimensions of information that are not objectively or directly given by the situation (for example, as instructions or as direct outputs from the environment), the more we can assume the presence of higher and more flexible levels of integration" (p 27). The reverse should also be true. That is, the more complex the perception the more it should contain dimensions of information not objectively or directly given by the situation. This indirect information must come from an inference process. As a result, we would expect that an inference process plays a greater role in the formation of perceptions about complex products than simple products. If people have employed an inference process more readily in the evaluation of complex products than simple products, we would

expect, based on learning, that people would more readily make inferences from indirect information about complex than simple products.

It is necessary, in order to test this proposition, to distinguish between products in terms of their complexity. Cognitive complexity has been measured in a number of different ways (c.f., Vannoy, 1965). One of the most frequently employed techniques has been Kelly's Role Construct Repertory Test (Rep Test). People are asked to provide the ways in which known stimuli (i.e., parent, teacher, child) are similar to and different from one another. People who provide many differences in response to this task are considered more cognitively complex regarding this stimulus domain than those who provide only few differences. Bieri, et. al. (1966) modified the Rep Test by providing bi-polar adjective scales rather than having the subject employ his own construct dimensions. Their results with the modified test have been consistently positively correlated with the results obtained from Kelly's original instrument. This modified Rep Test will be employed here to measure product complexity.

Product Related Cues

If we are interested in determining the effects of cues on product perception we must employ cues which do not, by their very presence, bring to mind previously formed, well-defined product perceptions. In studying cue utilization we are interested in determining the effects of cues on the formation of product perceptions. Use of these types of cues will permit us to explicate the nature of the process by which cues are combined to give meaning to the goods and services in the consumer's environment. Two cues which seem to meet these criteria are price and advertising. While the former has traditionally been employed in studies of product perception the latter cue is new.

Price

It has generally been found that, when price is the only cue, the subject perceives the higher priced product as being higher in quality (i.e., Tull, Boring and Gonsior, 1964; McConnell, 1968). When price is combined with other cues an interaction is frequently obtained (i.e., Andrews and Valenzi 1971; Jacoby, Olson and Haddock, 1971; Valenzi and Andrews, 1971).

Advertising

Nelson (1974) suggests that one of the most important pieces of information the consumer takes from advertising is merely the understanding that the brand is advertised. He contends, "the consumer believes that the more a brand advertises the more likely it is to be a better buy" (p 732). That is, he suggests consumers make inferences about a brand from knowledge of the magnitude of advertising support afforded that brand. The magnitude of advertising support given a product may be viewed, therefore, as a cue which permits inferences about that product without directly conveying product beliefs.

In the natural environment this belief that one brand is more heavily advertised than another is the result of exposure to many more advertisements for one brand than another. Nelson's contentions about the effects of this varied exposure could be subjected to experimental investigation by providing subjects in one condition with many advertisements for a brand while providing those of another condition with only a few or no advertisements. If actual advertisements were provided, however, the cue properties of advertising magnitude would be confounded with specific beliefs presented in the advertisement. One way to avoid this problem would be to give people

information about the magnitude of advertising support without actually showing them the advertisement. This is the procedure which will be followed here.

Theoretical Orientation

It has been maintained that product perceptions are formed by cue processing in accordance with a general inference process. Therefore, theories which account for the conditions under which and the extent to which people make inferences should be relevant to the prediction of product perception. Attribution theory is appropriate to this spectrum. While it has been used primarily to account for inferences made to other individuals, it is equally applicable to inanimate objects. That is, a person should employ the same process when making inferences about objects that he uses when attributing characteristics to other individuals. One of the most basic principles emerging from work in attribution theory is that attributions are directly related to the consistency of the meaning of cues associated with that object (c.f., Kelley, 1967). A stronger inference will be made when the attributional implications of the cues are consistent than when they are inconsistent. If high price is combined with high advertising support we would expect a stronger inference than if it is combined with low advertising or if only one cue is provided. We would expect then that a stronger, more extreme attribution will be made when cues are combined consistently than when they are combined inconsistently or when only one cue is provided.

Dependent Variables

If we view price and advertising as cues that people use to attach meaning to the products in their environment, it seems that dependent variables employed in the assessment of these products should give the consumer an opportunity to provide this meaning.

Osgood and his associates have found across a large group of studies, that the meaning of an object can be accounted for by the first three factors derived from a factor analysis of responses to a set of bi-polar adjective scales (i.e., Osgood, Suci and Tannenbaum, 1957). The object's position in a space comprised of these three dimensions is considered to be the object's meaning. The intensity of a person's position on a dimension is indexed by the distance of his position on that dimension from the midpoint. It is suggested, therefore, that the individual's position will be more extreme when cues are combined consistently than when they are combined inconsistently or when only one cue is provided.

Method

A pilot study was performed in order to select products differing significantly in complexity. This was followed by a laboratory experiment in which the implications of the previously suggested hypotheses were tested. In both studies the subjects were females recruited from civic organizations in the San Francisco Bay area. They volunteered in return for a donation made in their name to the civic organization or group from which they were recruited. The donations were \$1.50 each for those participating in the pilot study and \$3.00 per person for those taking part in the experiment.

Initial Study

Questionnaires were administered in a group setting to 61 subjects. Each subject received an instrument designed to measure cognitive complexity. Brands of black and white television, headache remedy, toothpaste and soft drink were used as stimuli. The brands of each product

class were placed on rows in a separate rating matrix. The columns of the matrices were occupied by bi-polar adjective pairs. The matrices provided the vehicle for measuring product complexity. They were developed in accordance with the procedure set out by Bieri, et. al. (1966). The order in which matrices were presented was randomized.

Experiment

Subjects arrived in the laboratory at the time of their pre-designated appointments in groups of up to eight people. After a general introduction by the experimenter, they proceeded to their cubicles. All treatments and questionnaires were administered individually on laboratory terminals while subjects were in these cubicles.

Subjects. One-hundred forty-four female volunteers participated in the experiment. Of this total, 118 subjects completed the task. The remainder failed to complete the experiment due to computer breakdowns and other malfunctions.

Independent Variables. Three levels of price (low, high and no-information control) were combined with three levels of advertising support (low, high and no-information control) and four products (black and white television, toothpaste, soft drink and headache remedy). Each subject evaluated each product under one of the price-by-advertising conditions. The order in which products were presented was counterbalanced.

Subjects were told that a new brand (in each of the given product classes) was about to be introduced to the local market. Variations in price and advertising support associated with this brand constituted the manipulations of the independent variables.

Price was manipulated by providing subjects with information about the suggested retail price of the new brand. Subjects in the high condition were presented with a price which was found to be higher than most brands in the given product class. Subjects in the low price condition were presented with a price which was found to be lower than that charged for most brands in the given product class. The control condition was given no price information.

In the low advertising condition it was stated that the brand will be "only lightly advertised." It was pointed out that advertisements will appear occasionally in local newspapers. In the high advertising condition it was pointed out that the brand will be "heavily advertised" and that it will appear frequently on "national network television" and in "national news and women's magazines". The control condition was given no information about advertising.

Product was manipulated by providing subjects with four different product classes. The products employed were chosen to represent both high and low complexity.

Dependent variables. After receiving a product description subjects were asked to evaluate the product on a set of 7-point bi-polar adjective scales. This was followed by a question which asked how likely they would be to buy the product if they were in the market for it. Finally, they were asked to provide their feelings about how high or low the product's price was and how lightly or heavily it was advertised.

This procedure was repeated three times (once for each remaining product). Subjects were then asked to respond to a series of post-experimental questions. These included a number of probes for demand characteristics.

After leaving their cubicles the subjects gathered in a

second waiting area. Their reactions were elicited. The purpose of the experiment was explained to them and any misperceptions were corrected. They were thanked for their participation and asked not to divulge anything about what went on during the experiment.

Results

Manipulation Checks

Complexity. The data from the pilot study were scored for cognitive complexity in accordance with the procedure developed by Bieri, et. al. (1966). Cognitive complexity is measured by comparing each rating in a row with the rating adjacent to it (i.e., for the same brand) in the other rows of the matrix. A score of one is given for every exact agreement of ratings on any one brand. This matching procedure is carried out in the same way for all possible pairwise comparisons in the matrix. The scores for each comparison are totaled to yield a final complexity score on the product class for each individual. Given an 8x13 matrix, the maximum, or simplest score obtainable would be $(13(13-1)/2)8$ or 624.

A cognitive complexity score was calculated in this manner for each subject and each product class. Orthogonal comparisons were made across products using multiple t-ratios. TV ($\bar{x}=140.3$), and soft drink ($\bar{x}=159.5$) were found to be significantly ($p < 0.001$) more complex than headache remedy ($\bar{x}=242.3$) and toothpaste ($\bar{x}=208.8$).

Advertising. After evaluating a given brand, subjects were asked how heavily advertised they believed the brand would be relative to other brands in the given product class. Responses were recorded on a 9-point equal appearing interval scale which varied from "Very Heavily Advertised" at the high end to "Very Lightly Advertised" at the low end. An analysis of variance was performed for each product class on the responses to this question. The manipulation of advertising was found to be significant ($p < 0.01$) in the expected direction for toothpaste, soft drink and headache remedy. It was found to approach significance ($p < 0.10$) for television.

Price. Subjects were asked how high in price they believed the test brand would be relative to other brands in the given product class. Responses were recorded on a 9-point equal appearing interval scale with the endpoints labeled "Very High" and "Very Low" respectively. Analysis of variance was performed for each product class on the results of this question. The manipulation was found to be significant ($p < 0.01$) in the expected direction for each of the four product classes.

Experiment

A principal components analysis was performed on a within-cell correlation matrix of bi-polar adjective scale scores. This adjusts the correlations to remove treatment effects. The components analysis was performed separately for each product class with varimax rotation of the first three components. Items loading more heavily than 0.500 on a given component were taken as representing that perceptual dimension. Items representing each of the first three factors for each product are shown with their loadings in Table 1. They represent the meaning or perception a la Osgood, Suci and Tannenbaum of each given product.

Individuals' scores on each factor were derived by summing their scores on items representing each factor. The items were scored +3 to -3, but, due to the fact that the earlier hypotheses concerned intensity or strength of perception, absolute values were summed to provide the dependent scores.

A separate 3x3 (advertising by price factorial multivar-

TABLE 1
Items Representing First 3 Factors For Each Product

Product	Factor 1	Factor 2	Factor 3
Television	Popular ¹ (.52311)	Superior (.75228)	Expensive to Buy (.60847)
	Modern (.66230)	Strong (.72662)	Expensive to Maintain (.66815)
	Good (.70565)	Complex (-.51814)	
	High Status (.65832)		
	Beautiful (.65124)		
	Sharp (.71704)		
Headache Remedy	Troublefree (.57121)		
	Good (.72049)	Modern (.61846)	Expensive (.64485)
	Lasting (.83351)	Safe (.62712)	Complex (.81128)
	Superior (.82179)	Popular (.52661)	
	Strong (.58562)	Speedy (.54344)	
Soft Drink		Healthful (.71764)	
	Carbonated (-.64400)	Easy to Digest (.74465)	
	Healthful (.85087)	Good (.70139)	Active (.62493)
	Natural (.79609)	Popular (.71338)	Expensive (.61316)
	Easy to Digest (.62565)	Superior (.54059)	Strong (.72435)
Toothpaste		Light (.62114)	
	Safe (.57458)	Good (.86918)	Strong (.71943)
	Smooth (.76374)	Healthful (.74441)	Complex (.80582)
	Pleasant Taste (.75496)	Superior (.70050)	
	Refreshing (.65650)		
	Creamy (.76317)		

¹Items represent the positive end of given bi-polar adjective scales. The numbers in parentheses are loadings.

TABLE 2
SUMMARY TABLE: Multivariate Analyses of Variance¹

Product	Effect	Multivariate F	p less than	Factors	Univariate F	p less than
Soft Drink	Price	2.278	0.038	Factor 1	0.437	.647
				Factor 2	1.269	.286
				Factor 3	5.191	.007
Television	Price	2.152	0.047	Factor 1	0.225	.799
				Factor 2	0.597	.553
				Factor 3	5.750	.004

¹Only significant effects are shown.

iate analysis of variance was performed for each product class. The dependent variables were individuals' scores on each of the three factors. A significant ($p < 0.01$) price main effect was obtained for television and for soft drink (see Table 2). No other significant main or interaction effects were obtained. Univariate analyses show that in both cases the price effect was due to factor 3. Multiple comparison tests (i.e., Tukey's Test, Kirk, 1968) indicated that the effects were due to the difference between the control and high price conditions. This can also be seen by referring to the table of means (Table 3).

Discussion

These preliminary results provide some support for the contention that cue utilization in product evaluation is directly related to product complexity. Price was related to product perception only for the two products found to be relatively complex. Further work is necessary, however, to substantiate this finding.

The results failed to support Nelson's contention that the amount of advertising provided a product is a cue to the product's value. The advertising manipulation check was found to be significant for three of the four products. Even though subjects differed significantly in their belief about how heavily the product was to be advertised, this difference did not affect their perception of the product.

The failure to find a significant cue effect for advertising prevented the testing of the hypothesis suggested earlier that the strength of the inference is directly related to the consistency in meaning of provided cues.

It would be desirable to employ a multi-method approach to the measurement of product complexity. The results provided by the Bieri, et. al. method could be compared with those obtained when other methods are employed. This could provide evidence to support the convergent

TABLE 3
Marginal Means For Significant Effects

Product	Price Treatment: Factor 3		
	Control Price	Low Price	High Price
Soft Drink	3.15	3.24	4.29
Television	2.23	2.85	3.33

validity of an instrument to measure product complexity. Vannoy (1965) found some evidence to support the convergent validity of instruments used to measure cognitive complexity. Work is in progress comparing results of the Bierl, et. al. approach with those obtained through factor analysis.

Further assessment of differences in cue utilization across product complexity levels would also be desirable. This could be accomplished by assessing the effects of cues on subjects' evaluations across a larger sample of products found to vary in terms of product complexity. If supportive results were obtained in these other studies it would make an alternative hypothesis that the results are due to product differences other than the typical complexity of its representation in cognitive structure less likely than it is from an assessment of only four products.

The failure to obtain stronger results in the study may be due, in part, to use of non-student women as subjects. These women, many of whom were older did not seem as comfortable and familiar as are students with the environment and tasks to which they were subjected. Most studies of price effects have employed students as subjects (Olson, 1976). It is likely, however, that different groups of people differ in their use of cues. It is also likely that complexity of a product's representation in cognitive structure differs by group. It is, therefore, important in this type of study to employ subjects representative of the target market for the chosen products. The women employed in this study seem to fit this criterion.

Olson (1976) has maintained that a theoretical orientation to future product perception research is necessary if we are to gain an understanding of cue utilization in product perception. It is believed here that attribution theory can contribute to this goal. Our consideration of the implications of cue consistency represents an initial basic hypothesis related to attribution theory. Many other hypotheses are also derivable from this work. In future research the theory could be adapted in a much more complete fashion. This could contribute greatly to our understanding of how, when and with what effect people employ cues in product evaluation.

Kelley (1967, 1973) employs four criteria in his treatment of attributional validity:

1. Distinctiveness: the impression is attributed to the thing if it uniquely occurs when the thing is present and does not occur in its absence.
2. Consistency over time: each time the thing is present, the individual's reaction must be the same or nearly so.
3. Consistency over modality: his reaction must be consistent even though his mode of interaction with the thing varies.
4. Consensus: attributes of external origin are experienced the same way by all observers. (Kelley, 1967, p 197).

Subjects in attribution studies are typically given a behavior sequence such as: John is the only person who laughs at the clown (i.e., McArthur, 1972). Subjects are then asked to attribute John's behavior to something about John (a person attribution), something about the clown (an entity attribution) or something about the situation (circumstance). When consensus (hardly anyone else laughs at the comedian) and distinctiveness (John also laughs at most other comedians) are low and consistency is high (John has almost always laughed at the comedian) an attribution will be made to the person. When consensus (almost everyone laughs at the comedian), distinctiveness (John does not laugh at almost any other

comedian) and consistency are high an attribution will be made to the entity (McArthur, 1972).

We may view an entity as a product (or brand). A product attribution will then be most likely when consensus, consistency and distinctiveness are high. Similarly we may expect the strongest product attribution when consensus, consistency and distinctiveness are high. If everybody buys brand B (high consensus); if John, for instance, does not buy any other brand (high distinctiveness); and if John always buys brand B (high consistency) then a strong attribution should be made to brand B. Presumably, the attributor would conclude that brand B is the best brand to buy.

We may adapt the theory even further by viewing consistency (in meaning) of cues related to the product as consistency over modality. That is, if every cue associated with the product provides a consistent implication then we have an instance of high consistency over modality. We may also wish to consider product class heterogeneity to represent distinctiveness. If high consistency and high distinctiveness are combined with high consensus we would expect a strong attribution to the product.

Further research would, of course, be necessary to substantiate these propositions. If people make inferences about products the way they do about other people then attribution theory should be relevant to the prediction of these inferences. Research directed at providing support for the propositions made here should contribute to a greater understanding of how and when people use cues in their evaluation of products. It should permit the marketer to do a better job of consciously and appropriately employing cues.

References

- I. R. Andrews and E. R. Valenzi, "The Relationship Between Price and Blind-Rated Quality for Margarine and Butter," Journal of Marketing Research, 7(August, 1970), 393-395.
- M. Bierl, A. L. Atkins, S. Briar, R. L. Leaman, H. Miller and T. Tripodi, Clinical and Social Judgment: the Discrimination of Behavioral Information (New York: Wiley, 1966).
- R. W. Gardner and R. A. Schoen, "Differentiation and Abstraction in Concept Formation," Psychological Monographs, 76(Whole No. 560, 1962).
- J. Jacoby, M. C. Olson and R. A. Haddock, "Price, Brand Name, and Product Composition Characteristics as Determinants of Perceived Quality," Journal of Applied Psychology, 55(1971), 570-579.
- H. H. Kelley, "Attribution Theory in Social Psychology," Nebraska Symposium on Motivation, 15 (1967), 192-238.
- H. H. Kelley, "The Process of Causal Attribution," American Psychologist, 28 (1973), 107-128.
- G. A. Kelly, A Theory of Personality: the Psychology of Personal Constructs (New York: Norton, 1955).
- R. E. Kirk, Experimental Design: Procedures for the Behavioral Sciences (Belmont, California: Brooks/Cole), 1968.
- L. A. McArthur, "The How and What of Why: Some Determinants and Consequences of Causal Attribution," Journal of Personality and Social Psychology, 22(1972), 171-193.
- M. D. McConnell, "The Price-Quality Relationship in an

Experimental Setting," Journal of Marketing Research, 5(1968), 300-303.

K. B. Monroe, "The Influence of Price Differences and Brand Familiarity on Brand Preferences," Journal of Consumer Research, 3(1976), 42-49.

P. Nelson, "Advertising as Information," Journal of Political Economy, 81(1974), 729-754.

J. C. Olson, "Price as an Informational Cue: Effects on Product Evaluations," Working Paper Number 43 (The Pennsylvania State University: College of Business Administration, 1976).

C. E. Osgood, G. J. Suci and P. H. Tannenbaum, The Measurement of Meaning (Urbana: University of Illinois Press, 1957).

H. M. Schroder, M. J. Driver and S. Streufert, Human Information Processing (New York: Holt, Rinehart and Winston, 1967).

E. M. Smith and C. L. Broome, "Experimental Determination of the Effect of Price and Market-Standing Information on Consumer Brand References." In R. Hass (ed.), Science, Technology and Marketing (Chicago: American Marketing Association, 1966), 520-531.

D. S. Tull, R. A. Boring and M. H. Gonsior, "A Note on the Relationship of Price and Imputed Quality," Journal of Business, 37 (1964), 186-191.

E. R. Valenzi and I. R. Andrews, "Effect of Price Information on Product Quality Ratings," Journal of Applied Psychology, 55(1971), 87-91.

J. S. Vannoy, "Generality of Cognitive Complexity-Simplicity as a Personality Constructs," Journal of Personality and Social Psychology, 2(1965), 385-396.

FAD FOOD USE AMONG ANGLO- AND MEXICAN-AMERICANS:
AN EXAMPLE OF RESEARCH IN CONSUMER BEHAVIOR AND HOME ECONOMICS

Joel Saegert, The University of Texas at San Antonio
Eleanor Young, The University of Texas Health Science Center at San Antonio
Merry Mayne Saegert (student), The University of Texas Health Science Center at San Antonio

Abstract

To home economists, consumption of fad foods is a behavior inconsistent with nutrition research data. The dynamics of belief in food fad benefits have not been satisfactorily explained by demographic or psychographic studies. The present study compared fad food use in two American ethnic groups. Suggestions for interdisciplinary cooperation in the study of food faddism are discussed.

Introduction

Both the American public and their public service representatives are becoming increasingly interested in the nutritional consequences of the products they consume. While such an increase in nutrition interest is a long-sought goal of home economists, some of the nutritional concern has manifested itself in the form of what nutritionists refer to as "fad foods." These products are considered to be a serious problem by many authors in the nutrition area (e.g., Bruch, 1974). Thus, the problem of fad food consumption seems to be an obvious area for cooperation between researchers in consumer behavior and home economics.

The nutrition research community has long refuted claims made by fad food advocates, either by showing that such claims have not been confirmed empirically (e.g., Anderson and Reid, 1974) or by pointing to research that shows that some fad foods can actually be harmful (e.g., Herbert, 1974). From the standpoint of nutrition education, the problem has been that fad food beliefs are held with what appears to be a very high degree of intensity (by certain segments of the population) and also that such attitudes do not seem to be susceptible to conventional educational attempts at combating ignorance and misinformation.

In the area of product attitude formation, consumer researchers have amassed a sizable literature on the multitude of variables postulated to be related to consumer purchase decision-making processes. It seems reasonable therefore to consider the problem of fad food consumption in the light of the theory and methodology which have been developed to study the formation of attitudes about general consumer behavior.

The Problem of Food Faddism

Fad foods, to the nutritionist, include a wide range of products and techniques which are offered to the public through various marketing activities. Some of these include "organic," "natural" or "health" foods, megadoses of vitamin and/or mineral supplements, quick weight loss diets, and more exotic items such as hair restoratives and sexual potency enhancers. In nearly all cases, such products are viewed by the nutrition research community as being harmful or as having little or no value, other than placebo effects, to consumers; these views are based on a broad foundation of scientific research and are supported by Federal agencies such as the Department of Agriculture and the Food and Drug Administration (e.g., Fusillo, 1974).

Numerous speculations have been offered to explain the consumption of fad food products in spite of lack of empirical verification of their value. Early hypotheses suggested that the fads were mainly used as "last resorts" by people who were desperately ill or aging and who had not responded to conventional medical treatment (Bruch, 1974; Jalso, Burns and Rivers, 1965). Recently however, researchers have argued that the demographic and psychological characteristics of fad food users "hardly confirm a stereotype of the rigid, stubborn and irrational 'health food nut'" (Calvert and Calvert, 1975). Typically, food faddists have described themselves as young, at least high school educated and in good or excellent health (Rhee and Stubbs, 1976). At least two studies have indicated that faddists cannot be classified as having bizarre or dogmatic personality characteristics (Calvert and Calvert, 1975; Saegert and Saegert, 1976). Thus, to both nutrition and consumer behavior scholars alike, the apparent paradox of an irrational purchase behavior being maintained by people with normal, middle-American background characteristics presents a challenging enigma for study.

An Example of Consumer Behavior/Home Economics Cooperation in Fad Food Research

One possible hypothesis concerning fad food use is that it is in some way associated with a middle-American value system which embraces such beliefs as the need for preventive maintenance and the utility of technological innovation to solve potential problems such as aging or future illness. If so, fad food use can be predicted not to be as prominent in cultural groups which stereotypically do not identify strongly with American middle-class values. For example, it is possible that Mexican-Americans may not subscribe to fad food beliefs in the same degree as Anglo-Americans of equal education and socioeconomic level.

Rhee and Stubbs (1976), in a study of 600 heads of households in two Texas cities found that Mexican-Americans did not differ from Euro-Americans in their use of fad foods. However, the random telephone sample used in their study contained only 3% (N=18) Mexican-Americans. To further study the hypothesis that Mexican-Americans do not use fad foods as frequently as Anglo-Americans, a larger sample was studied in another large southwestern metropolis which has a large (53%) Mexican-American population.

Method

Sample

Approximately equal numbers of Mexican- and Anglo-American shoppers were surveyed at a large shopping mall complex. The mall was chosen on the basis of its location near sizable middle socio-economic level neighborhoods of both ethnic groups. In all, 400 shoppers were interviewed; ethnic origin was determined by the interviewers as each questionnaire was administered.

Sampling was done on two weekends in the spring of 1977. Half of the interviewers were members of an undergraduate class in marketing research and half were members of an undergraduate home economics class in nutrition. A total of 16 interviewers participated.

Questionnaire

Shoppers were asked to fill out a questionnaire which contained items concerning several areas; among these were questions about fad product usage, general nutrition practices and demographic classification variables. Six items on the questionnaire were intended to provide information about usage of example products and purchasing practices which are typically associated with food faddism (Bruch, 1974): vitamins C, E, and B complex, "organic" foods, foods purchased at "health" food stores and "other" food supplements (examples were given as lecithin, minerals, etc.).

Earlier studies (e.g., Food and Drug Administration, 1972) have indicated that fad food habits are generally practiced without medical consultation. In the present survey, several items were included to insure that the food behaviors indicated by the respondents were not intended to correct diagnosed metabolic disorders. (Most of the interviewees, 72%, said that their use of the various products had not been prescribed by a physician.) It was also clearly stated on the questionnaire that vitamin usage should include only those products taken over and above "multi-vitamins."

The shoppers were approached in a systematic random manner (every "Nth" person to pass a specified point was approached). Approximately 10% of both ethnic groups declined to be interviewed. Usage rates of the six example fad foods and practices were measured by a four-level classification; the levels were "no use," "occasional use," "weekly use," and "daily use."

Results

The percentage of respondents in each response category for the two ethnic groups is presented in Table 1. As can be seen, significant differences ($p < .05$) were found for use of vitamins C and E; the difference in use of B vitamins approached significance ($p = .08$). In all three cases, the direction of the differences was as predicted; i.e., product use was less frequently reported by Mexican-Americans. Significant differences were not observed for use of organic foods, health food stores, or other food supplements.

Thus, slight differences in usage rates were observed for three of six example fad food practices considered in our survey. However, we also observed that there were some small but significant differences between the two ethnic groups on demographic variables, especially age, income and education level. These differences were such that the Anglo sample reported themselves slightly older, more educated and of higher income level (all p 's $< .02$ by Chi square test). In order to control for these differences, we felt that it would be useful to combine the product usage rates for the six fad items to provide a single overall index of food faddism. This score could then be subjected to an analysis of covariance which would compare ethnic groups using income, education level, and age as covariates.

Because of the low frequency of the "weekly use" response alternative, this category was combined with the "occasional use" category to yield a three-point scale. "No use" was assigned a value of zero, "occasional or weekly use" was assigned a value of one and "daily use" was assigned a value of two. The six scores for each respondent were summed to provide a scale which varied from zero (no use of any fad items) to twelve (daily use of all six products or practices). A slight difference in mean total food faddism score between ethnic

TABLE 1

Incidence of Fad Food Practices in Two Ethnic Groups

		No use	Percentage Responding			χ^2 (df=3)	p
			Occa- sional use	Weekly use	Daily use		
Vitamin C	Anglo-Americans	59	19	2	20	8.63	.03
	Mexican-Americans	60	26	3	10		
Vitamin E	Anglo-Americans	75	12	1	13	8.34	.04
	Mexican-Americans	76	12	4	8		
B Complex	Anglo-Americans	68	16	2	15	6.77	.08
	Mexican-Americans	76	12	4	8		
Other Food supplements	Anglo-Americans	82	10	2	7	5.31	.15
	Mexican-Americans	82	10	5	3		
Organic foods	Anglo-Americans	58	33	4	6	3.24	.36
	Mexican-Americans	50	37	7	8		
Health food stores	Anglo-Americans	74	22	2	2	5.44	.14
	Mexican-Americans	76	15	4	4		

groups (Anglo-American $\bar{X}=2.51$; Mexican-American $\bar{X}=2.38$) was found to be non-significant in the analysis of covariance ($F<1.0$). Thus, it must be concluded that overall fad food usage, as reported by the respondents in our survey, was equivalent for the two ethnic groups, even though a few differences were noted on three specific vitamins (before age, income and education were adjusted). It is not known whether this indicates that the two groups were equivalent in subscribing to middle-American values or that there is simply no relationship between "value system" and fad food practices (as hypothesized above).

Discussion

The present survey data fail to provide evidence of a difference in fad food use among two American ethnic groups. This demonstration of the null hypothesis is not startling when considered along with other research in the fad food area. Relatively few empirical comparisons have been made between fad food users and non-users, but those which are available have failed to isolate specific reasons for subscription to fad food beliefs. Moreover, there is no reason to suspect that fad food users are other than healthy, youthful, affluent, educated, normal Americans; the present study merely adds ethnic origin to the list of factors which cannot be used to explain the phenomenon of fad food use.

Such results are also not surprising to students of consumer behavior. Although it is intuitively obvious that demographic and/or psychographic variables are related to purchase behaviors and product attitudes, the consumer behavior literature is notoriously full of reports of failures to pinpoint simple relationships (e.g. Frank, Massey and Lodahl, 1969).

Rather than dismiss the, as yet, unfathomable phenomenon of fad food use by the American public as insuperable, the area can serve as an example of the need for combined effort through inter-disciplinary cooperation. The following suggestions are some areas in which such interdisciplinary efforts may be useful.

Increased use of Empirical Research and Statistical Analysis

While the literature in home economics and nutrition has provided a wealth of theoretical speculations on the dynamics of fad food use (Bruch, 1974; McBean and Speckman, 1974; Schafer and Yetley, 1975), relatively few empirical studies have been undertaken. Also, some of those which have been conducted have not taken advantage of analytical techniques which might tease out explanatory relationships. For example, an important study by Calvert and Calvert (1975) used a standardized measure of the socio-psychological variable of dogmatism to investigate the degree to which fad food users were characterized by irrational belief systems. Unfortunately, the study did not compare food faddists to a control group of non-fad food users nor were the percentages of users holding rational vs. dogmatic beliefs in various categories tested for statistical significance. This study would have been substantially more useful if more rigorous procedures of research design and analysis had been used.

Increased use of Modern Multivariate Analytical Techniques

Marketing theorists have often pondered the relatively low correlations found between personality traits (as measured by standardized tests) and usage of specific products. The answer must lie in the fact that such relationships are too complex to be found through

simple one-to-one correlational techniques. In what is now a classic study on personality and product use, Sparks and Tucker (1971) used canonical analysis, a technique which considers complex relationships between sets of dependent variables and sets of independent variables. They found that relationships of relatively high magnitude emerged when patterns of personality traits were compared with patterns of product usage behavior. For example, the trait "sociability" was correlated with one set of products in one personality trait milieu, but was correlated with an entirely different set of products in the context of a different milieu.

Such complex relationships obviously exist in the case of fad food usage. The employment of such multivariate techniques as canonical analysis, factor analysis, discriminant analysis and multidimensional scaling should reveal underlying relationships which can be used to understand food faddism.

Increased use of Consumer Information Processing Models

Recent theory in the area of consumer behavior has seen the development of cognitive information-processing models to try to integrate the complex network of factors which must underlie consumer decision-making. Such models have taken diverse forms; some examples include Howard's recent assessment of consumer behavior from the standpoint of cognitive psychology theory (1977), Bettman's decision-net model (1974) and Kuehn's linear learning model of brand choice (1962). Variables often included in such models are attention and memory, advertising media influence, susceptibility to novel ideas, psychographic factors, and economic market conditions.

Fad food behavior can provide interesting data for consideration in such models. For example, advertising processes for fad food products seem to be different from those of other products. Since the benefits claimed for fad products are not substantiated by empirical evidence, touting such benefits in advertisements is precluded by false advertising laws. Therefore, promotions in ads for fad foods only provide notice of their availability and price, rather than of their proposed benefits or advantages. Claims of powers for such products often appear in articles in health food periodicals but it is left to the consumer to make the connections between the benefits advocated in the articles and the advertisements for the products. Understanding how such processes fit into models of advertising effectiveness would make an interesting contribution to both the consumer behavior and nutrition education areas.

Conclusion

It is suggested that the phenomenon of fad food use is merely an example of a much broader problem in American society. This problem is one in which the "establishment" sources of information, i.e., scientific researchers, government agencies and the food industry, maintain one position with respect to available evidence, while a sizable group of consumers have chosen to ignore or deny that view. Such behavior is observed in beliefs ranging from the benefits of Laetrile, the safety of food preservatives, and the benefits of acupuncture to such entertaining but widely discussed beliefs as unidentified flying objects, pyramid power and the Bermuda Triangle. The issue involves the overall relationship between consumer attitude-formation and traditional, research-based information. Cooperation among researchers in the fields of home economics and consumer behavior can provide insight into the mechanisms behind such conflict between data and belief. However, we should obviously not expect substantial

breakthroughs in understanding this problem without a great deal of concerted collaboration among researchers in consumer behavior, home economics and a host of other disciplines.

References

- T. W. Anderson, and D. B. W. Reid, A double blind trial of vitamin E in Angina Pectoris. The American Journal of Clinical Nutrition, 27 (1974) 1174-1178.
- J. R. Bettman, Decision-net models of buyer information processing and choice: Findings, problems and prospects. In G. D. Hughes and M. L. Ray (Eds.), Buyer/Consumer Information Processing. (Chapel Hill: The University of North Carolina Press, 1974, 59-74).
- H. Bruch, The allure of food cults and nutrition quackery. Nutrition Reviews, 32 (1974), 62-66.
- G. P. Calvert, and S. W. Calvert, Intellectual convictions of "health" food consumers. Journal of Nutrition Education 7, (1975) 95-98.
- Food and Drug Administration. A study of health practices and opinions. U.S. Department of Health, Education, and Welfare, PB 210 978, (June, 1972).
- R. D. Frank, W. F. Massy, and T. M. Lodahl, Purchasing behavior and personal attributes. Journal of Advertising Research, 9 (1969) 15-24.
- A. Fusillo, Food shoppers beliefs: Myths and realities, FDA Consumer, (October, 1974).
- V. Herbert, Destruction of vitamin B₁₂ by ascorbic acid. Journal of the American Medical Association, 230 (1974) 241.
- J. A. Howard, Consumer Behavior: Application of a Theory. (New York: McGraw-Hill Book Co. 1977).
- S. G. Jalso, M. N. Burns, and J. M. Rivers, Nutritional beliefs and practices. Journal of the American Dietetic Association, 47 (1965) 253.
- A. A. Kuehn, Consumer brand choice as a learning process. Journal of Advertising Research, 2 (1962) 10-17.
- L. D. McBean, and E. W. Speckman, Food faddism: A challenge for nutritionists and dieticians. American Journal of Clinical Nutrition, 27 (1974) 1071-1078.
- K. S. Rhee and A. C. Stubbs, Health food users in two Texas cities. Journal of the American Dietetic Association, 68 (1976) 542-545.
- J. Saegert and M. M. Saegert, Consumer attitudes and food faddism: the case of vitamin E. Journal of Consumer Affairs, 10 (1976) 156-169.
- R. Schafer and E. A. Yetley, Social psychology of food faddism. Journal of the American Dietetic Association, 66 (1975) 129-133.
- D. L. Sparks and W. T. Tucker, A multivariate analysis of personality and product use. Journal of Marketing Research, 8 (1971) 67-70.

THE PROBLEMS, PITFALLS, AND OPPORTUNITIES IN INTERDISCIPLINARY APPLIED CONSUMER RESEARCH

Joseph Barry Mason, The University of Alabama
Helen Goetz, The University of Alabama

Abstract

This paper explores the problems and opportunities inherent in the pursuit of interdisciplinary research and outlines various strategies for encouraging faculty involvement in such activities.

Introduction

The problems of interdisciplinary research are discussed infrequently in the published literature and few empirical studies have been published on the topic. Thus, few role models exist. One exception is the work being pursued at the University of Washington under a National Science Foundation grant as part of the Research Management Improvement Program (e.g., Birnbaum, 1975; Gillespie, 1976; Mason, 1975).

Further, the concept and practice of interdisciplinary research is fraught with myths and misconceptions. Also, few persons have invested the time and energy necessary to establish truly interdisciplinary research activities. Disagreement even exists on the meaning of the term "interdisciplinary." The terms interdisciplinary, multi-disciplinary, and transdisciplinary are often used interchangeably by colleagues discussing the subject. For the sake of clarity the following definitions are presented.

- Discipline..... A specific body of teachable knowledge with its own background of education, training, procedures, methods, and content areas.
- Multidisciplinary.. Juxtaposition of various disciplines, sometimes with no apparent connection between them. e.g.: music + mathematics + history.
- Pluridisciplinary.. Juxtaposition of disciplines assumed to be more or less related. e.g.: mathematics + physics, or French + Latin + Greek: "classical humanities" in France.
- Interdisciplinary.. An adjective describing the interaction among two or more different disciplines. This interaction may range from simple communication of ideas to the mutual integration of organising concepts, methodology, procedures, epistemology, terminology, data and organisation of research and education in a fairly large field. An interdisciplinary group consists of persons trained in different fields of knowledge (disciplines) with different concepts, methods, and data and terms organised into a common effort on a common problem with continuous intercommunication among the participants from the different disciplines.
- Transdisciplinary.. Establishing a common system of axioms for a set of disciplines, e.g., anthropology considered as "the science of man and his accomplishments." (Centre for Educational Research and Innovation,

1972, pp. 25, 26).

Most persons probably agree that the major problems of contemporary society are increasingly interdisciplinary in nature, but neither interdisciplinary research nor education are simple processes, especially given the discipline-oriented university and the mission oriented society of today. Ideally, interdisciplinary research efforts involve pooling the talents of several faculty members from various disciplines who integrate their knowledge in conducting research on a problem that individual researchers working alone could not effectively resolve. Too often, however, multi-disciplinary research is the end product, not interdisciplinary research. As noted, "multi-disciplinary research... can be performed by experts with interdisciplinary backgrounds, but who work separately, not necessarily in the same environment or within a mutual confrontation, exploring different aspects of a central problem. Results of multi-disciplinary research may be integrated by one of the researchers or by someone else (Nilles, 1976, p. 80; also, Segner, 1973).

The lack of ready acceptance of interdisciplinary research efforts is at least partially due to the nature of the university. Disciplinary structures are the basis of the university and of the professions engaged in teaching and research. The arrangement provides a convenient breakdown of knowledge into readily identifiable parts. However, disciplines as we know them today do not represent a preordained order of knowledge. The dynamics of social phenomena forces interchange between disciplines and even creates new disciplines. "The inter-discipline of today is the 'discipline' of tomorrow" (Centre for Educational Research and Innovation, 1972, p. 9).

Unfortunately, most persons experienced in interdisciplinary research agree with the statement that "a principle feature of interdisciplinary research is not necessarily profound scientific development, but an improvement in human relations to overcome the intolerance, bigotry, and barriers not only among individuals, but within and among disciplines" (Mar, 1976, p. 65; also, Sherif and Sherif, 1969; Piaget, 1970).

The discussion that follows reflects our joint experience in interdisciplinary research, graduate and undergraduate teaching, development of interdisciplinary curricula, and the direction of interdisciplinary masters theses and doctoral dissertations. The thoughts presented also reflect the experiences of colleagues at other institutions with whom we have discussed this area.

Potential Problems in Interdisciplinary Research

Colleague support of interdisciplinary research activities. The quality and worth of much interdisciplinary research is suspect in the eyes of many colleagues. This often occurs because of the blending of concepts from two or more disciplines in an interdisciplinary research activity, but typically not the use of frontier concepts from either discipline. Thus, the research activity is likely to reflect breadth of knowledge, perhaps more than depth of knowledge. Faculty members are trained in doctoral programs and selected for faculty appointments on the basis of their knowledge

of and research skills appropriate to a particular discipline, not because of interdisciplinary skills. Thus, suspicion of the unknown is to be expected.

We have clearly seen the suspicion with which interdisciplinary research is viewed by various professional associations. For example, a session focusing on consumer education at the secondary school level was recently proposed to the program committee of a major national consumer association with an avowed interdisciplinary thrust. The proposed session, including three empirically based papers, was not accepted because the program committee felt that the topic was not appropriate for sponsorship by the organization. Likewise, when two of our marketing colleagues presented a paper at the American Council for Consumer Interests meeting a few years ago and approached the problem from a marketing perspective--development of marketing strategies for reaching consumers--instead of from the home economics perspective--consumer welfare--criticisms of the paper and the approach were both intense and vocal. Much progress in the recognition of the value of interdisciplinary research remains to be accomplished.

Recognition and support by administrators. Faculty members who pursue research activities outside their own primary area of expertise face the possibility that this will be viewed as contrary to the goals of their college and to the goals of the individual faculty member, defined as achieving excellence in a particular discipline. For example, we know of a young home economics colleague who, with the support of the department chairman, began to pursue interdisciplinary research with young faculty members in the marketing department of a college of business. The chairman of the marketing department also supported this effort. The result yielded several quality articles which were published in such journals as the Journal of Social Psychology. However, the Dean of the College of Home Economics indicated to the young faculty member that this activity was largely a waste of time as it did not contribute to the visibility of the College of Home Economics or contribute to the conceptual and technical skills of the individual in terms of solving problems unique to home economics. As part of the process of chastisement, support to the individual in the form of research monies and computer time were lessened. Also, it was made clear the research accomplishments in no way contributed to progress toward promotion and tenure.

Finally, one further example illustrates the point. One of the authors recently collaborated with a young faculty member in writing an article from the young person's doctoral dissertation. The logical publication outlet for the article appeared to be the new British journal, the Journal of Consumer Studies and Home Economics. However, the young faculty member indicated that he preferred not to have the article published in this journal because the publication would be viewed by his colleagues and his dean as a relatively insignificant contribution which would not be acceptable by a marketing journal. These narrow viewpoints and outright prejudices against interdisciplinary research can be very disheartening.

In summary, interdisciplinary research in a university setting often faces major human relations problems and requires a continuing education of other faculty members and administrators on the value of such activity. In the absence of support by peers and administrators, the path is particularly fraught with difficulties for a young faculty member in a promotion and tenure track. However, the benefits to the individual and society can be tremendous, and these comments are not designed to frighten people away from interdisciplinary research efforts but simply are offered to point out some of the pragmatic realities which often exist today.

Ways of Generating Interdisciplinary Research

The following discussion reflects on the various activities in which we have engaged over the past ten years and which have been useful in developing and encouraging an interdisciplinary research environment in our departments. The various approaches are relatively easy to implement and can yield significant benefits.

Strive for better communication. Communication is probably the fundamental problem in interdisciplinary research. The persons involved must understand the different viewpoints focusing upon a central research problem. Without this appreciation, which requires a high level of tolerance and patience, good interdisciplinary research simply cannot occur. Establishing adequate communication is a time consuming process even under the best of circumstances and may well take several years. Otherwise, what happens is a multi-disciplinary effort.

Each discipline has its own sophisticated jargon which can only be understood over time. To many people, this process of establishing constructive communication may seem scientifically unproductive. We have observed that persons who are strongly committed to a narrow academic specialization have difficulty in participating in interdisciplinary research for this very reason.

Further, it is necessary to recognize that disciplines often have intolerance for other disciplines. Representatives of each discipline are likely to claim superiority or intellectual distinction over the others. For example, we probably all have heard a colleague ask the question, "What can a home economist possibly tell me about consumer behavior?" or "Those faculty members in marketing are simply interested in finding ways to help businesses exploit the consumer; they have no interest in consumer welfare."

How, then, can communication be encouraged? The following possibilities seem to work fairly well.

Cross-listed courses. One of our major breakthroughs occurred when Professor Joseph Uhl, an agricultural economist at Purdue University, taught a graduate consumer behavior course at the University of Alabama which was cross-listed between home economics and marketing. Students from both disciplines enrolled in the course. The resulting dialogue was lively and instructive. At least one doctoral dissertation emerged as a result of the discussions in the course, and a heightened awareness of and interest in interdisciplinary problems emerged on the part of both students and faculty members.

Interdisciplinary curricula. In 1970, the consumer sciences curriculum in the College of Home Economics was revised at both the graduate and undergraduate levels to require or encourage consumer science students to take courses outside the College of Home Economics. Even though our joint interests in interdisciplinary activities extend back to 1968, the curriculum revision which occurred accelerated the interest and placed several graduate and undergraduate home economics students in the marketing curricula on a regular basis.

Colloquia. Several seminars for graduate students and faculty have been conducted in recent years in which research ideas and research findings have been presented. Also, focuses on common methodological concerns have been a part of these discussions. Lastly, areas of common consumer research interests are likely to emerge.

Joint appointments. Joint appointments, if more than simply an arrangement on paper, can be especially useful in encouraging an appreciation of the values of interdisciplinary research. Such arrangements allow

badly needed dialogue at faculty meetings, give a measure of credibility to the activities of the faculty member which would otherwise be lacking, and encourage broadened dialogue with graduate students and faculty members who share joint interests in a given area of research.

Masters theses and doctoral dissertations. Two interdisciplinary doctoral dissertations have been completed under our joint direction. The focus of one was on evaluating the effectiveness of consumer education at the secondary school level, and the second dealt with consumer cognized distance and the factors which affect distance perception. Five or six masters theses on diverse aspects of consumer behavior have also been completed under our supervision in recent years.

Students and faculty members have to be selected carefully for participation in this process. For example, students must be capable of tolerating a higher level of ambiguity than is normal because of the complexities of working with faculty members from more than one discipline. Also, we have found that the senior faculty members involved have to be constantly on the alert so that the student is not "used" by a faculty member from one of the disciplines in generating research which is primarily disciplinary in nature, but which is conducted under the guise of interdisciplinary research. Intellectual snobbery often occurs when a young faculty member makes disparaging remarks about the other discipline involved.

Interdisciplinary doctoral programs. We have been successful in structuring one doctoral level curriculum and one masters program in the last few years which have been truly interdisciplinary. These were housed in the Graduate School of the University to avoid the problem of departmental barriers. These vehicles are useful in allowing students to select courses from a wide array of disciplines such as marketing, economics, consumer sciences, psychology, geography, sociology, and other areas which are appropriate to a particular area of study. The result is a stronger conceptual and methodological base for conducting research.

Interdisciplinary research committees. The two of us have made a continuing effort to serve on committees which are evaluating possible areas for interdisciplinary research. These have included activities related to the Center for the Aging, Center for the Study of Drug and Alcoholism Abuse, the New College of the University, the Health Care Management Advisory Committee and similar programs or institutes. These arrangements provide a unique opportunity for advocating quality interdisciplinary efforts.

Advantages of Interdisciplinary Research

Development of a broader philosophy. Inevitably, interdisciplinary research activities will lead to a greater understanding and appreciation of the philosophies underlying the respective disciplines of marketing and consumer science. The horizons of the participants are expanded. Problems are more readily viewed from a variety of different perspectives and a greater tolerance for diverse viewpoints occurs. Also, entirely new dimensions of rather conventional research problems emerge. As recently noted, for example, marketers often treat the "consumer not as the subject, but as an object to be affected in order to realize their behavior goals" (Arndt, 1977, p. 13). Interdisciplinary research helps overcome this bias.

Language barriers are overcome. As a result of interdisciplinary efforts, faculty members will benefit by becoming able to read a greater variety of journals.

Language barriers are resolved and levels of heightened intellectual awareness and appreciation occur.

Development of new conceptual frameworks. The value of differing conceptual frameworks for focusing on a given problem is quickly realized. Alternative assumptions and viewpoints emerge which make one more sensitive to the nuances of language and of the need for developing shared meanings to permit linking of insights into the disciplines involved. An interdisciplinary focus also fosters the convergence of the two disciplines in developing more comprehensive consumer interest and public policy perspectives and helps conceptualize public policy on matters of broad social concern.

Current Research Activities

This paper is not primarily designed to report on our ongoing research. However, a brief discussion of how the research originated and its focus will help to reveal some of the everyday realities of interdisciplinary research. The focus of our current research effort is the senior citizen and the marketplace. Funding was provided by the Center for Aging on the University of Alabama campus. It quickly became apparent to us that an interdisciplinary focus on senior citizens shopping behavior was badly needed. Thus, the specific areas of focus of the research include the shopping behavior of the senior citizen, information seeking activities, food consumption patterns, and satisfaction and dissatisfaction with various products or services and complaint behavior, nutritional adequacy of diets, and the adequacy of physical facilities within the senior citizen household.

Logical Areas for Interdisciplinary Research

Numerous opportunities exist for meaningful interdisciplinary research between the Colleges of Business and Colleges of Home Economics. A few areas quickly come to mind and include various dimensions of public policy, metrication, consumer education, quality of life, the handicapped, use of leisure time, improving consumer competence, and family resource use. For example, a recent conference at the University of Nevada identified the following areas of research in the generally defined area of "improving consumer competence in family resource use":

- A. Comparison shopping studies for both goods and services.
- B. Family expenditure for food, clothing, shelter, services.
- C. Assessment of non-food items charged to the food budget.
- D. Comparison of convenience foods and/or prepared foods with those made at home.
- E. Evaluation of information available to consumers in relation to potential purchases.
- F. Comparison of costs in different retail outlets.
- G. Assessment of knowledge regarding consumer credit.
- H. Attitudes and values of different groups about money and financial management.
- I. Knowledge of credit costs and available information about credit.

Summary

Truly, the opportunities for and the values of interdisciplinary research are many. However, persons contemplating such activities should be fully aware of the potential pitfalls and frustrations and the large amount of time necessary to establish communication links which are necessary for an atmosphere of mutual trust and support. Finally, one needs to be prepared to answer questions about the relevance and quality of the research from persons in both disciplines.

However, the potential problems are well worth the benefit, particularly for persons who are well established in their respective disciplines and who can afford to pursue interdisciplinary research even in the face of some uncertainties and risk. As noted, "working outside the conventional structure is costly in time, energy and general strain. Its only conceivable justification is that it opens up educational territories that could otherwise remain unexplored" (Southern Regional Education Board, 1972, p. 6).

Really, an interdisciplinary focus is a mental outlook and a desire for self fulfillment through different approaches to problems. It necessitates openmindedness and requires continuing practice and flexible mental patterns. Properly approached, it can be a new stage in the development of scientific knowledge as is reflected in the papers presented in the recently released Association for Consumer Research publication entitled The Broadened Concept of Consumer Behavior (Zaltman and Sternthal, 1975).

Strong support and recognition of the value of interdisciplinary research, such as shown by the development of this forum, also is a vital and necessary force for effectively overcoming disciplinary barriers and rigidities. Clearly, peer recognition at the national level of the value of such activities is a fundamental requisite, given the realities of promotion and tenure guidelines in most major universities today.

References

- Johan Arndt, "A Critique of Marketing and The Broadened Marketing Concepts," paper presented at the Second Macro-Theory Seminar, University of Colorado, Boulder, August 1977.
- Phillip H. Birnbaum, "Management of Interdisciplinary Research Projects in Academic Institutions," unpublished doctoral dissertation, University of Washington, 1975.
- Centre for Educational Research and Innovation, Interdisciplinarity: Problems of Teaching and Research in Universities (Organisation for Economic Co-operation and Development, 1972).
- D. F. Gillespie, "Research Team Effectiveness in the Academic Setting," unpublished doctoral dissertation, University of Washington, 1976.
- Brian W. Mar, William T. Newell, and Borje O. Saxberg, "Interdisciplinary Research in the University Setting," Environmental Science and Technology, 10 (1976).
- James Roy Mason, "Management of Organized Interdisciplinary Research Units in Universities: Selected Cases," unpublished doctoral dissertation, University of Washington, 1976.
- Jack N. Nilles, "Interdisciplinary Policy Research and the Universities," IEE Transactions on Engineering Management, Vol. EM-23 (May, 1976).

Jean Piaget, "General Problems of Interdisciplinary Research and Common Mechanisms," in Main Trends in the Social and Human Sciences (Paris: UNESCO, 1970).

E. P. Segner, "Administrative Structure for Interdisciplinary Research: A Panacea or a Can of Worms?" A paper presented at the Workshop on Implementation of Interdisciplinary Research, 81st Conference of the American Society for Engineering Education, Ohio State University (June 29, 1973).

Muzafer Sherif and Carolyn W. Sherif (eds.) Interdisciplinary Relationships in the Social Sciences (Chicago: Aldine Press, 1969).

Southern Regional Education Board, Regional Spotlight: Interdisciplinary Explorations in the South (Atlanta: Southern Regional Education Board, 1972).

Gerald Zaltman and Brian Sternthal (eds.) The Broadened Concept of Consumer Behavior (Chicago: Association for Consumer Research, 1975).

SOURCE SIMILARITY AND FASHION NEWNESS
AS DETERMINANTS OF CONSUMER INNOVATION

Karen Kaigler-Evans, The University of Texas at Austin
Clark Leavitt, The Ohio State University
Lois Dickey, The Ohio State University

Abstract

Using a 4 x 4 x 4 fractional replication, the research combined 4 different fashion ideas, in each of 4 levels of newness, and four pictured sources rated from highly similar to dissimilar to the population from which the 48 subjects in the experiment were chosen. For each subject one idea from each newness level was attributed to each source immediately before the subject rated her agreement with the idea. The main effects of both source similarity and fashion newness were significant as hypothesized.

Introduction

The diffusion of innovations and the adoption or non-adoption of new products and ideas is an area that has been cultivated by several disciplines including home economics, rural sociology and marketing. The related topic of social influence has occupied the attention of sociologists particularly. Most of the research studies in these areas have been field studies that have structure variables that underly diffusion such as social strata and homogeneity of membership groups or the role of key individuals in relaying information and influence: innovators and opinion leaders, for example.

At the same time social psychologists also have been interested in the adoption of new ideas under the guise of attitude change and persuasion. However, the research paradigm, unlike that of home economics, has included the laboratory rather than a field setting and has been concerned largely with the individual processes underlying the adoption of new attitudes and beliefs. Social psychologists, together with students of mass communication, have been interested in the role of mass media in social influence and have dealt in this context with two general classes of independent variables, those related to the source and to the message.

There has been remarkably little cross-fertilization between these two research traditions of innovation and attitude change. The home economists' concern with innovation or education is essentially an application problem in need of conceptual enrichment while attitude change research has been characterized by a great deal of theory, but is unseasoned by serious attempts at application. The research reported here is an attempt to combine elements of both traditions. It uses the laboratory paradigm of attitude change research while attempting to enrich the independent variables in a way most appropriate to the research tradition familiar to the home economist, particularly that of social influence and fashion adoption. Primarily we have attempted to meld variables from both disciplines and have taken over only the necessary minimum of conceptualization from either field. Before describing the experiment, we shall review the ideas from each field that might interact profitably with those from the other.

Social Influence and Fashion Adoption

Three kinds of studies may be mentioned briefly: the psychological significance of clothing and fashion; population trait distribution related to innovators and opinion leaders, and the relation of social structure

to adoption behavior. Each provides evidence for the relevance of the variables used in this research.

Studies in the literature of home economics deal with the function of clothing as both a cue used by others to make social judgments and as a means of expressing the wearers own internal needs or values (Gibbons, 1969; Compton, 1962; Douty, 1963; Rosencranz, 1962; Connor, et al., 1975). This research is only indirectly relevant here: it makes clear that clothing is rich in personal meaning and that it can also be ambiguous as to what it can mean for the self.

There has been a sizeable body of research on social influence both in relation to fashion and to other products and ideas. Informal influences or word-of-mouth activities have long been of interest to social scientists. The two-step flow of influence proposed by Lazarsfeld and Merton, (1954), is a classical attempt to explain the lack of an immediate, obvious effect of mass media by invoking informal networks as leveraging agents. Most of this research has attempted to identify the kind of people who are influential in adoption and who adopt early in the life cycle of an innovation. In particular, King, (1963), has used such research to challenge the traditional concept of fashion as a vehicle for status differentiation that "trickles down" from higher to lower strata. Instead he emphasized the importance of peers or similar others as sources of fashion information (see Grindering, 1967).

More recently a different set of constructs has been invoked to examine the spread of ideas in informal communication. These are variables applied to communication networks, in particular the interconnectedness and the homogeneity of an individual's network. A person's network of social contacts can vary from homogeneous interconnectedness to a state of more heterogeneous, looser connections. At some point between these two extremes, conditions are right for maximum influence.

Any particular person is more likely to be exposed to new ideas when interacting with a person who is different, but at some point the difference will become so great that communication will break down or fall off. The balance point of greatest effectiveness between these competing tendencies has been called the point of optimal heterophily, a term originally introduced by Merton, (1957) and made current by Rogers, (1976). Similar concepts underlie other field studies. Laumann, (1973), found that men in interlocking networks were less open-minded with respect to non-conformity than those in radial networks. Interlocking networks are more homogeneous in respect to religious and occupational status while the radial networks represent weaker ties and greater heterogeneity. The idea that we learn more or are similar to us has been the subject of other field studies. Granovetter, (1973), and Liu and Duff, (1972), have attempted to demonstrate what they called the "strength of weak ties" for the diffusion of innovations. People are more likely to accept suggestions from others with whom they have only weak ties.

Studies of Source and Message in Attitude Change

Much of the research mentioned above has been concerned with specific products or fashions. Most studies also

are concerned with the characteristics of the person who influences others. These two variables are paralleled by message and source determinants in persuasion studies. We shall not review message research here except to point out that innovations and fashions are characterized by newness which is a frequently used variable in psychology, both in terms of familiarity and novelty seeking. It is often regarded as an underlying determinant of cognitive arousal.

Source variables have been approached in two ways that are relevant here. First, basic perceptual dimensions of the source of communication have been studied as interpersonal perceptions. Typically three dimension appear: expertise or competence, trustworthiness or safety, and attraction or dynamism (Giffin, 1967). (A similar approach has been used with message variables in which case novel and useful seem to emerge as basic dimensions.)

The second source variable of relevance here is similarity of the source to the receiver. Two contrasting interpretations have been made regarding the role of similarity. Byrne and others have demonstrated a relation between similarity and attraction which lends credibility to the "myth of super-representativeness" (Simons, 1973). This refers to the popularly held notion that influential speakers are seen as being similar to the audience -- "one of us." Clearly, however, an expert source would not be similar to the receiver -- the more expert, the less similar. Alpert and Anderson, (1973), found a curvilinear relation between similarity and influence -- not surprising in light of these considerations.

Whichever the direction of relation in any particular case, the variables of similarity of source and of newness of the idea are highly relevant to the problem of describing the underlying dynamics of consumer innovation. This is especially the case in view of the recent trend toward examining social influence and communication in terms of peer groups and homogeneous networks. The experimental procedure followed here not only used these variables but attempted to operationalize them in as representative a manner as possible in a laboratory setting.

In an experiment done to find the point of "optimum heterophily," Alpert and Anderson, (1973) used 3 students as presumed sources of statements about air travel. They were presented in photographs and varied from normal appearing to deviant and were shown to student subjects drawn from the same population (business administration). The pictured source that was most influential (evoked most agreement with statements attributed to him) was the one that was intermediate in magnitude of similarity to the subjects as measured by profile similarity on a series of rated personal traits.

The present study extends these findings to fashion suggestions and also varies the degree of newness of the suggestions. More important, it defines similarity in a more general way using self/similarity ratings of a group of independent judges to preset the source values. To the extent that this method is judged to be successful the results raise the question of whether similarity to the self is the critical variable or whether it is an artifact of group representativeness stemming from group similarity ratings.

We hypothesize that the relation between source similarity and influence or agreement with suggestions made by the source will be curvilinear with middle level sources being most influential (the so-called inverted U curve) as would be predicted by the notion of optimum heterophily.

Method

While the experiment to be reported here was straightforward, several auxiliary studies were done beforehand to prepare the experimental materials and some additional measures were gathered on the subjects themselves in a later session. We shall first describe the preparation for each of the two experimental manipulations: the source of the innovations and the message about the innovation itself.

Selection of Persons for Sources

In an effort to maintain some of the naturalism of field studies, we used as sources of the fashion suggestions snapshots of actual coeds taken on the campus of a large midwestern university where the experiment itself was to be carried out. The photographs had a casual look as though taken impulsively by a friend, although the size, background and camera angles were relatively uniform. Forty photographs were taken in as unsystematic a way as possible using different times and places and types of subjects.

From this pool, sixteen pictures were selected that seemed to represent the range of women typical of home economics undergraduates. A group of 63 home economics undergraduates, none of whom knew the women in the snapshots, were asked to judge each photographed woman on a five point scale, ranging from "extremely similar to you" to "not similar at all." Using the method of summated ratings, the sixteen photographs were scaled and four were chosen covering equal intervals of the scale (3.53, 2.85, 2.42 and 1.98 were the actual values of the mean similarity ratings).

The scale represents group similarity because the self ratings of similarity to each person in the group, when combined, become an index of the similarity of the photographed person to the entire group. (We felt that a direct rating of similarity to the group would be a meaningless task.)

Selecting of Fashion Suggestions

Although it would have been possible to bring actual fashion innovations into the laboratory, such a strong manipulation would have overwhelmed the source variable. In any case, influence is most often manifest in talking about innovations before the actual trial. To maintain realism in representing the fashion object we asked women from undergraduate classes to write down fashion suggestions that they actually had heard recently. The term "suggestion" was not given a limited interpretation but included all forms in which comments might be made such as "I like ...". A total of 78 women contributed 352 items. After editing and eliminating duplication there were 58 acceptable fashion comments.

Next we asked a panel of five buyers from local stores catering to students to rate each statement for the extent to which it was a new idea for local college women. They sorted fashion comments into four groups from extremely fashionable to not-so fashionable. The four levels of fashion newness included such statements as "Long dresses are fun to wear to parties" and "Shoes with rope trim are comfortable." Sixteen statements were selected, four from each newness level, for presentation to the subjects in the experimental treatments. All but one statement lacked unanimous agreement concerning the newness level.

Subjects

The main experiment was carried out with 48 undergraduate women enrolled in home economics. None were used in any of the previous steps. Each was assigned to one of 4 groups in a Latin square design pairing each of the 4

items at each newness level with one of the 4 photographed sources using a different combination in each of the 4 groups. Thus each of the sixteen items was paired with each source an equal number of times, the order being counterbalanced in the Latin square design. The photographed source was shown using a 35 mm slide projected on a screen in front of the subjects.

When the picture of the source was projected on the screen the subject indicated on a five-point scale her agreement with the fashion opinions attributed to that source. Each subject made judgments about 16 items, four from each source.

A second session of the experiment required all subjects to rate each source on a series of trait descriptors. In this session, following the main experiment by one week, each subject was given a personality scale to measure open versus cautious information processing (Leavitt and Walton, 1975), a measure of affiliation or liking for each source (Hayes, 1972), and the series of adjective ratings to describe the sources (reported in detail elsewhere).

Results

Complex analysis of variance was used to determine the significance of the results using Plan 12 of Winer (1962, page 571). This is a 4 x 4 x 4 balanced fractional replication Latin square in which the treatments are assumed to be fixed and subjects within groups are a random variable. The three main effects are the four sources, the four newness levels and the four replicated fashion suggestions within each newness level (assumed to be identical in the design of the study). The results are contained in Table 1.

The main effects of sources and newness level were significant. Their interaction was not significant for all groups combined, but when the groups are considered separately they have a significant interaction. Replicates by groups interaction is also significant indicating that response to the individual items was determined not only by newness level and source pairing but also by their unique qualities. This failure to achieve uniform response to items within newness levels -- while not surprising -- makes the interactions in the study difficult to interpret.

The mean agreement level for the four sources is shown separately for each newness level in Figure 1. Even though the first order interaction is not significant, combining the newness levels may be misleading. It can be seen that the effect of similarity is most unequivocal for the two extreme newness levels while the two middle levels show the inverted U curve found in previous research.

The overall effect of the group similarity of the four sources is curvilinear in roughly the same manner as the two middle levels of newness. The mean agreement from most to least similar was 3.6, 3.4, 3.7, and 2.9. The exception to the trend was source II. In order to investigate this further, the scores were plotted separately for the thirteen subjects who scored highest and the eleven subjects who scored lowest on the open processing scale given in the post session. This is a measure of the extent to which the person adopts an open, constructive attitude toward innovation. The results are graphed in Figure 2. Clearly it is source II that distinguishes the two types of processors.

To explore this difference further, the average rating for all four sources on the 26 trait ratings made by all subjects in the post session was computed. The strong and weak traits for each source are listed in Table 2. It can be seen that source II and source III are almost

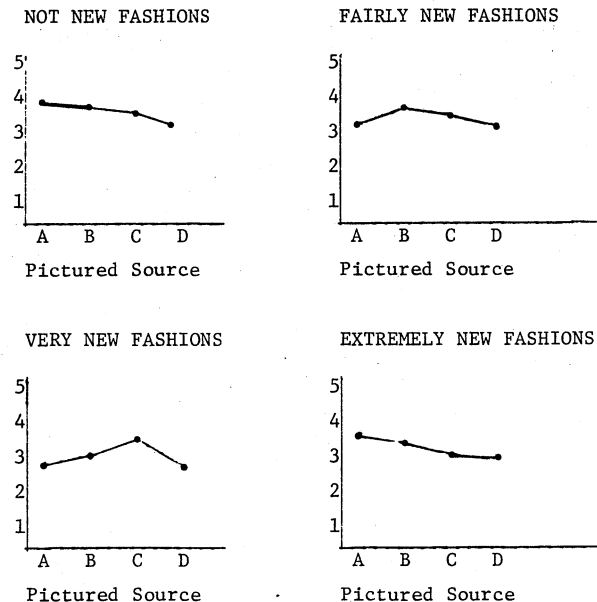
mirror images. Source II is high on anxious and aggressive and low on conventional and trustworthy while source III is the reverse except for conventional.

TABLE 1

ANOVA: AGREEMENT WITH FASHION SUGGESTIONS OF FOUR NEWNESS LEVELS ATTRIBUTED TO FOUR PICTURED SOURCES

Source of Variation	SS	df	MS	F	P
<u>Between Subjects</u>	<u>72.480</u>	<u>3</u>			
Group	.886	3	.295		NS
Subjects w. groups [error a]	71.594	44	1.627		
<u>I Within Subjects</u>	<u>1022.509</u>	<u>720</u>			
B (Sources)	16.074	3	5.358	3.544	.05
B x Group	22.895	9	2.544	1.683	NS
B x Subjects w. groups [error b]	199.531	132	1.512		
C (Newness)	57.803	3	19.268	13.617	.001
C x Group	8.458	9	.940		NS
C x Subjects w. groups [error c]	186.739	132	1.415		
BC	9.478	9	1.053		NS
BC x Group	99.728	27	3.694	3.469	.01
BC x Subjects w. groups [error bc]	421.803	396	1.665		
<u>II B x Group</u>	<u>22.895</u>	<u>9</u>			
A (Replicates)	14.865	3	4.955	1.901	NS
AB'	8.030	6	1.338		
BC x Group	<u>99.728</u>	<u>27</u>			
AC	69.765	9	7.752	6.654	.001
AB'C	29.963	18	1.165		

FIGURE 1



AGREEMENT OF SUBJECTS WITH FASHION INFORMATION VARYING IN NEWNESS ACCREDITED TO FOUR PICTURED SOURCES (Source A is most similar and 5 is agree extremely well)

This is interpreted as indicating that open subjects are less dependent on conventional credibility cues such as expertise and more likely to respond to a source who is somewhat controversial as a source per se, but more representative of the group with which the subject identifies.

FIGURE 2

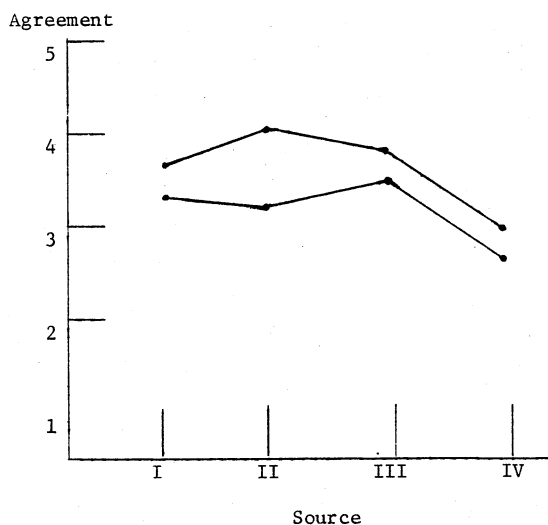


TABLE 2

PERCEIVED PERSONAL TRAITS OF THE SOURCES

Source	Strong Traits	Weak Traits
Source I	active sophisticated innovative aggressive	anxious conventional introverted
Source II	anxious aggressive	conventional trustworthy
Source III	trustworthy expert introverted	anxious aggressive
Source IV	anxious conventional introverted	active sophisticated innovative expert

Conclusion

What do the results tell us about the psychological processes underlying consumer innovation? They clearly demonstrate the importance both of characteristics of the source and of the innovation itself for acceptance of fashion ideas. The attempt to ground laboratory experimentation in the real world by using naturalistic manipulations -- snapshots and "overheard" fashion ideas -- was only partially successful since the realistic flavor of the fashion items seemed to increase the "noise level" of the analysis: the four replicates within each newness level contributed significant interaction variance in the design. This could have been avoided by constructing

sterile phrases designed to produce response only in terms of the one aspect of interest -- newness. Also, the effects of the photographs were more difficult to interpret than would they have been if only one dimension were varied -- hair color for instance. But, all-in-all, these individualistic variations give us more faith in the generalizability of the results than would have been the case if a less naturalistic stimulus variation had been used.

Our measure of similarity was more generalizable than Alpert and Anderson's, but the operation that makes it so -- independent judgments -- raises the question of whether the essential variable is similarity of source to the individual receiver or to the group as a whole (super-representativeness). Further research will be required to unconfound these two interpretations.

References

Mark Alpert and W. Thomas Anderson, "Optimal Heterophily and Communication Effectiveness: Some Empirical Findings," Journal of Communication, XXIII(September, 1973), 328-343.

Norma Compton, "Personal Attributes and Color Design Preferences in Clothing Fabrics," Journal of Psychology, LIV(1962), 191-195.

Barbara Connor, K. Peters and M. Nagasawa, "Person and Costume: Effects on the Formation of First Impressions," Home Economics Research Journal, IV(September, 1975), 32-41.

Helen Douty, "Influence of Clothing on Perception of Persons," Journal of Home Economics, LV(1963), 197-220.

Keith Gibbins, "Communication Aspects of Women's Clothes and Their Relationship to Fashionability," British Journal of Social and Clinical Psychology, VII(December, 1969), 301-312.

N. Giffin, "The Contributions of Studies of Source Credibility to a Theory of Interpersonal Trust in the Communication Process," Psychological Bulletin, 68(1967), 104-119.

M. Granovetter, "The Strength of Weak Ties," American Journal of Sociology, 78(May, 1973), 1360-80.

Margaret Grindering, "Fashion Diffusion," Journal of Home Economics, Vol. 59, No. 3, (March, 1967), 171-174.

Charles W. King, "Fashion Adoption: A Rebuttal of the 'Trickle Down' Theory," in Stephen A. Greyser, ed., Toward Scientific Marketing (Chicago: American Marketing Association, 1963), 108-125.

Edward O. Laumann, Bonds of Pluralism: The Form and Substance of Urban Social Networks (New York: John Wiley and Sons, 1973).

Paul F. Lazarsfeld and Robert K. Merton, "Friendship as Social Process: A Substantive and Methodological Analysis," in Monroe Berger, Theodore Abel and Charles H. Page, ed., Freedom and Control in Modern Society (New York: D. Van Nostrand Co., Inc., 1954).

Clark Leavitt and Karen Kaigler-Evans, "Mere Similarity Versus Information Processing: An Investigation of Source and Message Interaction," Journal of Communication, 2(1975), 300-6.

Clark Leavitt and John Walton, "Development of Scale for Innovativeness," in Mary Jane Schlinger, ed., Advances in Consumer Research, Vol III(1975), p. 545.

W. T. Liu and R. W. Duff, "The Strength in Weak Ties," Public Opinion Quarterly, 36(Fall, 1972), 361-66.

Robert K. Merton, Social Theory and Social Structure, (Glenco, Illinois: The Free Press, 1957).

Everett Rogers, "New Product Adoption and Diffusion," Journal of Consumer Research, No. 4(1976), 290-306.

Everett Rogers and Dilip K. Bhomik, "Homophily and Heterophily: Relational Concepts for Communication Research," Public Opinion Quarterly, XXXIV(Winter, 1970-1971), 523-538.

Mary Lou Rosencranz, "Clothing Symbolism," Journal of Home Economics, LIV(1962), 18-22.

Herbert W. Simons, "Perception and Communication," in C. David Harper and Row, Publishers, ed., Advances in Communication, (1973), 215-240.

B. J. Winer, Statistical Principles in Experimental Design (McGraw Hill, 1962).

CLOTHING DECISIONS: A DECISION PROCESS ANALYSIS OF FOCUSED GROUP INTERVIEWS

Roger D. Blackwell, The Ohio State University
Jo Ann Schickel Hilliker, The Ohio State University
(On leave from the Kentucky Cooperative Extension Service)

Abstract

The purpose of the research was to gain a comprehensive understanding of variables and the processes involved in the purchase of selected garments of women's clothing. A decision process approach was used to analyze ten focused group interviews conducted in varied geographic and demographic populations.

Introduction

The purchase and use of women's garments has been investigated by many researchers. Generally, these investigations have concentrated on only one or a few of the variables such as information sources (Stauffer, 1972; Arbaugh, 1974; Orsini, 1972; Martin, 1971-1972), confidence and perceived risk related to satisfaction (Sproles, 1969), physical characteristics such as color or fiber content (Wylie, Grown, and Morris, 1977; Schultz and Phillips, 1976) racial and socioeconomic factors (Braguglia and Rozencranz, 1968; Patson, 1971; Sturdivant, 1971; Schickel, 1970), social class (Warning, 1956; Roach 1960; Kundel, 1976), quality/price relationships (Stafford and Enis, 1969; Sims, 1969), and numerous other evaluative criteria such as laundering qualities, appearance and style, durability, fit and so forth (Borrell et al., 1964; Nolan and Levine, 1959; Shelly, Goldberg and Clayton, 1968; Clayton and Sherman, 1972; Jenkins and Dickey, 1976). Many studies have also focused on complaints and satisfaction (Whitlock et al., 1959; Conklyn, 1971; Steiniger and Dardis, 1971; Ayers et al., 1963; Wall, Dickey and Talarzyk, 1976).

The investigation of each of these variables provides important clues to understanding purchase and usage decisions but the clues are limited only to partial explanations. A more comprehensive approach is needed to integrate all of these variables, or even to determine what variables are relevant. Ryan (1966, p. 178) has commented on the need for understanding a wide range of variables and their interrelationships:

We can have a better understanding of what will be satisfactory to the consumer if we know: why people choose the clothes they do; how society influences them in their selection; the relationships between personal values, interests, attitudes, self-concepts and personality factors and the effect of clothing on individuals.

The research reported in this paper takes an integrated approach to understanding clothing decisions, using a comprehensive model of consumer decision processes as an aid to collecting and analyzing data. This study involves qualitative data and must be regarded as exploratory in nature. However, it is based on an approach designed to discover major variables important in clothing decisions which may later be subjected to more quantitative approaches. The results, though exploratory, do describe how some consumers make clothing decisions and may therefore be useful both to management interested in marketing clothing and to home economists interested in maximizing satisfaction from the use of clothing.

Research Methodology

This research involved focused group interviews analyzed by reference to the Engel-Kollat-Blackwell (EKB) model of decision processes. (Engel, Kollat, Blackwell, 1973). The results are reported in this paper, using the EKB model as a framework for presentation.

Data Collection

Ten focused groups were conducted in varying geographic regions. Each group consisted of six to twelve women (average of 10) selected by commercial marketing research organizations using random recruiting techniques and screening questions to obtain balanced groups. The screening question given to randomly selected women on the telephone consisted of a statement in which respondents were asked to rate themselves as "very interested, somewhat interested, or not interested" in fashion. Only those who were very interested in fashion were asked to participate. Other questions established the woman's age, dress size, and socio-economic characteristics. Each group consisted of a balanced composition of age and socio-economic characteristics and three racial groups. Thus, because of recruiting specifications in this study, the women wore average clothing sizes and possessed a greater than average fashion interest compared to the population as a whole.

The interviews were conducted in Atlanta, Boston, Columbus (Ohio), Daly City (California) and San Antonio. All interviews were tape recorded and conducted in specially equipped facilities of commercial marketing research organizations. Half of the interviews were conducted by the senior author and half by another investigator experienced in group interviews.

Focused Group Approach

The purpose of the focused group approach is to provide a sampling of the decision elements in clothing decisions. No attempt is made to develop a statistical description of the proportion of consumers in the population possessing specified attributes. The qualitative approach of focused group interviews does not supplant statistical description but rather provides an important prior step by discovering variables that are important and generating hypotheses for investigation in quantitative studies (Cox, Higginbotham, Burton, 1976).

An interview guide was used in conducting the focused group interviews. This generally enhances the efficiency of the technique in collecting the maximum amount of relevant information but at the same time has the potential of arresting discussion that might generate topics not previously considered by the researchers and may force discussion of topics that are not meaningful to the group participants (Merton, Fiske, Kendall, 1975). These problems were recognized by the interviewers who, being sensitive to the problems, attempted to allow and encourage the discussion into areas not covered in the interview guide (but relevant to fashion and clothing topics) and to move to other areas of discussion when topics failed to generate meaningfulness to the participants.

Theoretical Framework for Analysis

The Engel-Kollat-Blackwell (EKB) model of consumer

decision making was used for analysis of the data and as the theoretical framework for the study. For each garment analyzed in the study, separate analysis pages were prepared containing headings corresponding to stages of the EKB model (problem recognition, internal search, etc.). Each of the authors independently listened to all tapes of the interviews and recorded each comment relating to a stage in the EKB model. Other comments that did not "fit" in the model but were of interest in understanding fashion and clothing purchases were defined as follows:

Problem recognition is the consumer decision stage which is caused by recognition of unmet needs. Problem recognition results when a consumer recognizes a difference of sufficient magnitude between what is perceived as the desired state of affairs and what is perceived as the actual state of affairs. It leads to doing something about this feeling or recognition.

Internal search and alternative evaluation is the consumer's use of existing attitudes to identify and evaluate alternative solutions. The consumer identifies the pertinent evaluative criteria based on past experience.

External search and alternative evaluation is the consumer's use of various sources of external information such as mass media, personal sources and marketer-dominated sources (advertisements, dealer visits and so on) to learn about the number of alternative solutions to the perceived problem. The characteristics and attributes of these alternatives and the relative desirability of each is reviewed and weighted by the consumer.

Purchase process is the customer in-store environment interaction. Purchase process behavior can be precipitated by two events. One, it can be problem oriented in the sense that the consumer visits a retail outlet in order to purchase a product or service that satisfies some perceived problem. Second, other factors such as desire to get out of the house, a desire to get away from the spouse and children, a desire to engage in fantasy may motivate a visit to a retail store.

Outcome refers to consequences of a decision and the various forms of resulting behavior. Two outcomes are specified: (1) triggering of new behavior, and, (2) post purchase evaluation.

A second objective of the study involved understanding contemporary conceptions of fashion and understanding of the relationship between such conceptions and clothing purchases is necessary to predict future trends that may occur in fashion conceptions and clothing purchases.

The remaining portion of the paper describes the results from the focused group interviews. Following a general discussion on the concept of beauty and role of clothing, the results concerning decision stages are reported for each garment analyzed in the study. These garments include a good dress, skirts, casual clothes, pant suits, bras and panties. Due to the overlap between some garment categories, only the results which relate to distinguishing attributes or decision elements are reported for some garment categories.

Contemporary Conceptions of Beauty and Clothing

Contemporary women viewed beauty and fashion in diverse ways and the only consensus position was that a beautiful woman is one who looks healthy, clean, well-postured and appears to take care of herself. Among young women, this usually meant a "natural" look but among middle-aged women, there was more tolerance for "some help for nature."

The women in this study consistently defined beauty more as "personality" or the inner qualities of a person rather than external appearance. The quality of a relationship with a person or a person's attitudes toward life and other people were increasingly viewed as beauty -- a position that de-emphasized the role of clothing in achieving "beauty."

When pressed to describe women of physical beauty, two different models emerged. One was of the contemporary, natural women. She has a free spirit, is liberated, and wears both casual clothing, such as jeans and casual shirts, or wears exotic dresses with extreme cleavage, no bras and often back-less design. The other model of beauty was the woman who appeared healthy and had a good body. She wears attractive clothing but not exotic clothing.

The Role of Clothing

Clothing was worn to fit into women's peer groups or her life style. Clothing was much more individualistic in the sense that fashions were determined by many types of groups or segments rather than one fashion for the society as a whole. A typical comment reflected this view: "People notice clothing because it explains a lot about your personality." Another woman amplified this view when she said, "You can't judge a book by its cover but clothing gives a good clue about which people are alike."

A Good Dress

Respondents indicated that a good dress was not as important in their lifestyles as it once was. Pants have replaced the dress for many occasions.

Problem Recognition

Varied comments were made indicating the recognition of a need for a new dress. The purchase process for a dress was usually initiated by a special occasion or by a recognition of new fashion look. Career women, however, often walk through a store and see a dress which triggers the decision process.

Many women mentioned that they purchase a dress for a special occasion such as a wedding or funeral or to wear to a special place such as a nice restaurant. In general they felt a dress was expected or more appropriate than pants for such occasions.

Some comments seemed to indicate that fashion is one reason for recognizing the need for a new dress. Several women stated that in spring or fall they feel the need to look for new clothing. The recognition of this need appeared to be a combination of boredom with their present clothing and an awareness of the "new" fashions.

Several of the younger women indicated that they bought a dress because they did not have one. The recognition of a need for a dress may have been prompted by two influences. First, at the time the interviews were conducted (1975), the pants fashion had been important for several years. It had reached a peak and had begun a decline as the "fashion." Skirts and dresses were beginning to appear. Since the women interviewed possessed above average interest in fashion, they probably were picking up a trend. In other words, the eye was becoming accustomed to a new look. Secondly, several women said their purchase of a dress was influenced by husbands or boyfriends. A typical expression offered by one woman was, "My husband said, Why don't you buy a dress? I always see you in pants." Several respondents stated a belief that men were pleased to see legs. Their beliefs support Laver's theory of shifting erogenous zones in women's dress (Laver, 1964).

Career women reported that dresses were still appropriate for office wear, at least two or three days a week, or for special "dress up" days. High school students reported that dresses were worn by some to school but mostly by the straight kids who were "out of it" or by the very stylish, country club kids. There was considerable agreement that students who wore a dress all the time were considered "strange."

Internal Search

The criteria for evaluation of a dress varied with each woman. It appeared that most women went shopping with an idea of what they intend to buy but looked at several dresses to find the one that met most of their criteria. Whether the individual liked the dress when in the store was an important criterion. Color was also very important in the selection of a dress.

Women reported that they must try on quite a few dresses to find the right fit. Some stores were viewed as never having anything that fits or that is appropriate while others do.

Price was also important. Women usually had a price range in mind and did not deviate much from it. Some women made a point to browse through sale dresses.

The care of the dress was also a consideration. Some women would only consider buying washable garments; other preferred dry cleanable ones. One woman said that since she owned only one dress, she preferred to have it dry cleaned.

Several women expressed concern related to the construction of the garment. Several inspected the garment to determine if it was durable. They expected the dress to last a long time. Comments were also made concerning the thread, buttons and size of seams. The women most concerned with construction appeared to be homesewers.

External Search

Few women relied directly on mass media for information on dresses. A few girls indicated that they liked the clothes of Mary Tyler Moore and Cher on television. Most women consulted personal sources of information such as a sister, mother, boyfriend, or husband. High school girls indicated that the popular people at school were significant influences.

Purchase Process

The comments related to purchase process fell into two categories -- store related and product related.

Store related. The store considerations concerned its location, the sales personnel and whether or not an individual "likes" the store. Only a narrow range of stores appeared to be considered. The respondents desired a store that was easy for them to get to. Career women reported that they often do their shopping during lunch hour and choose a store convenient to work.

Various attitudes toward salespeople were expressed. Some women indicated a positive attitude toward sales women but most expressed a desire to look at the dresses and not be helped by salespeople. A typical comment was expressed by one young woman, "Clerks bug you. They always say that a dress looks so cute even when you know it looks crappy."

The product. Two types of dress shoppers were revealed in this study. One is the woman who likes to shop in large department stores for a good selection. The other is the woman who prefers small, unique shops. The latter type disliked shopping in department stores because

"you see yourself on the street."

SKIRTS

Most of the comments made concerning a good dress also applied to skirts. The primary attraction of skirts was the variety in skirt lengths, which made skirts a desirable change from pants.

There was a general concensus among the women of all ages that mini skirts were out of fashion even for younger girls. Respondents stated a belief that short skirts became accepted because of the youth revolution, liberalism, freedom and new sexual attitudes; however, they believed that now the mood is more conservative, resulting in more skirts of long and moderate lengths.

Several women indicated that for special occasions they selected the floor length or maxi skirts because they were "fun to wear" and made them feel "special". These women appeared to be very receptive to the fashion trend toward long skirts for "at home" wear or for entertaining.

The women were also concerned about the fit of skirts. Some tall girls reported that it was easier to find skirts to fit than pants that fit. Others indicated difficulty in finding skirts that fit around the waist and hips.

CASUAL CLOTHES

Casual clothes were defined by the researchers as clothes appropriate for home or leisure activities. Almost all women referred to jeans or pants with assorted tops as casual clothes. Other possible clothing items, such as culottes or shorts, were not included in the current definition of casual clothing by most women.

Jeans were recognized as the favorite of younger women although several women over 30 commented on their pleasure in wearing jeans, too. It appears that jeans will be included in wardrobes for many years in the future. Some typical comments of respondents illustrate the probable future in jeans:

"We'll still be wearing jeans twenty years from now."

"They are trying to get skirts back instead of jeans but it won't work, except maybe long skirts, because jeans are so comfortable."

"I like jeans because you can wear anything with them. It doesn't matter if you get a stain on them. You can just keep wearing them every day -- and they are accepted everywhere."

"I wear jeans around home 95% of the time," (Ex-school teacher, age 35).

"We'll never stop wearing jeans."

Problem Recognition

Pants and jeans were perceived to be ideal for contemporary lifestyles. They were worn almost everywhere and were described as the most "comfortable" garment. Almost all women owned at least one pair. The wardrobe of some younger women contained only pants or jeans; no dresses or skirts were included.

Pants and jeans were purchased when old ones wore out. As a woman's inventory became depleted, she became more receptive to purchasing jeans or pants in routine

shopping. The women were also sensitive to special sales and special displays or other stimulants to sales.

As with dresses, the purchase of pants or jeans was also prompted by a need related to an activity. The role for pants and jeans has expanded to include all activities including work, parties, dinners on the town and even religious activities. The style of pants or jeans varies with the activity.

Internal Search

Jeans were important to women and they were willing to expend considerable effort to find jeans or pants that match their evaluative criteria. There were specific features that women looked for, such as type of cuff, durability, fit.

Fit was the most important evaluative criteria. Women wanted jeans and pants that fit. A major problem for many women was to get jeans to fit around stomach. Some respondents reported loyalty to some brands because of fit but others said that even the same brands changed so often in fit that brand loyalty does not help. Women reported searching for jeans that fit and finding few that do.

Several women commented on the pros and cons of all cotton versus cotton blend jeans. All cotton jeans were liked because of the softness, however, they wore out quickly. Other women desired cotton blended jeans because of their durability.

A small but distinct group of women indicated that they liked jeans because they could embroider favorite designs on them. They particularly liked to express their own individuality in the designs chosen. This group tended to be younger women who disliked their clothing to be like everyone else.

External Search

Store displays were credited with influencing the type of casual clothes purchased. Husbands and family members were also important influences. An older woman explained that she had not worn pants for 50 years but that at the instigation of her 20 year old daughter she began wearing them for yard work. Now she likes them so much she wears them every place and stated that now they could not be taken away from her.

Purchase Process

Jeans are bought everywhere. Selection is an important factor in determining which store is shopped. Stores appealing to youth -- such as those on campuses -- were perceived to have better selection.

Quite a few women reported the need for shopping in several stores in order to find jeans or pants that fit. The respondents reported that they always tried on pants to check the fit. The exact fit desired depended on the individual. Some women wanted jeans not too tight; others wanted jeans that would shrink to fit their bodies. Some women complained that jeans were too tight in the hips and too loose at the waist, requiring something to be worn over the waist to cover it up. Several respondents stated the male jeans fit better because they could buy the correct leg and waist measurements.

PANT SUITS

The respondents indicated that the use of pant suits varied with age. The younger women felt that pant suits were worn only by older women. Several high school girls indicated that their mothers and grandmothers wore pant suits almost everywhere -- to work, parties, movies

and other places. Their fathers were credited with influencing the mothers to wear pant suits. The younger women emphasized that they preferred the mix and match quality of jeans or pants and tops rather than the coordinated suit. Pant suits were generally considered to be too "dressed-up" for younger women.

The older women interviewed indicated that they liked the freedom and comfort of pant suits. Pant suits were perceived as "right" for the office. An older woman explained that pant suits were for older women over 25 who wanted to be neat and who would not wear jeans but did not want to wear a dress. They were believed to be good for lifestyles involving either work or bridge in the afternoon.

BRAS

Attitudes toward bras have become more realistic and more individualistic. Respondents viewed a bra as something to be worn if a woman needs it rather than simply because every one else wears one or does not wear one. If a woman is small, she may not need one but a larger woman probably needs one, especially in active situations such as sports. A few years ago when some women stopped wearing bras, it was considered a "showy" thing with shirts or blouses that made it obvious that no bra was being worn. Today, it is less for show and more a matter of individual preference.

Problem Recognition

Buying a new bra usually occurred for one of two reasons. First the woman became increasingly aware the bras she owned did not provide support, had sagging straps, were discolored, or did not make her feel as she wanted to feel. If the situation was really serious the woman might make a special trip to a trip to a department specialty store. By the time she entered the store she had determined the need for a new bra. Thus, she was particularly susceptible to ads or special promotions in the store.

The second reason that often resulted in the purchase of a new bra was the purchase of outerwear that required a different type of bra. Sometimes the purchase of an especially nice dress caused feelings that it was wrong to wear an old bra with a special occasion dress. A new bra might be purchased at the same time as the dress.

Internal Search

The criteria with which women evaluated and selected bras were varied and reflected a highly segmented market. There was nearly complete disagreement or absolute polarity reflected by respondents' comments about what they wanted in a bra. Bra features considered included total disagreement about whether straps should be elastic (so they don't "cut" you and allow active movement) or non-elastic (so they don't stretch out of shape). Some women asked for wider backs because it was more comfortable; others preferred the opposite because wide straps interfered with current fashions. Some women preferred the seamless bras which gave a natural, soft, rounded appearance. The women who liked seamed bras indicated that these bras lasted better and gave better support in contrast to the seamless which became flimsy with a few washings.

The women agreed that the "pointy" look was undesirable and in decline. There was also general agreement that bras must be easily washable in automatic machines without discoloring or falling apart.

Prettiness was an attractive feature for bras but not sufficient reason to cause women to buy the bra or become a satisfied customer. The appearance of the bra

may have caused the woman to try it on or give it consideration but other criteria were used in making the purchase.

The two criteria used most often by women in the actual purchase of a bra were fit and comfort. Each woman may define these two criteria differently.

Women may look at the same bra and have very different evaluations on these criteria. Some women look at underwire support bras and say, "How awful. That would just kill me." Other women look at the same underwire bra and say, "That has a good underwire and looks like it would give me enough support to be comfortable."

External Search

As a result of the decision to purchase a bra made prior to entering a store, women often wait for sales or shop for special sales before purchasing. The need for a new bra may be evident but time is usually not crucial.

Younger women were influenced by their friends and their preferences in bras. The women indicated that as their figures matured, each woman relied more on her personal evaluative criteria.

Purchasing Process

Women bought bras in stores that offered a large selection. Middle-aged women were more brand loyal than young women. Young women were more likely to select in bargain basements, specialty stores, chain department stores (Sears, Penneys) and some discount stores. They looked at the prettiness, price, color, and so forth before selecting. The middle-aged woman bought the same brand and style purchased before if satisfied with it and it was available. Several women complained that the styles changed frequently and that they were unable to repurchase a favorite style.

Women felt that they must try on a bra before buying it. The exceptions were young women of moderate (usually B cup) size who were able to buy the one size soft bras and the D cup woman who had limited selection, and such precise requirements that she became a repeat purchaser of one style and brand.

Women were divided in their opinions about the helpfulness of sales clerks in the foundation departments of retail stores. Many women reported that few clerks were helpful with fit problems and style substitutions.

PANTIES

Panties were considered a "fun" product for most young women rather than a product to be taken seriously. Most women reported that they wore briefs and preferred a lot of variety in applied decoration and color.

Problem Recognition

Panties were purchased when the ones owned wore out. Store displays or sales frequently reminded the consumer that her inventory was low.

Several younger women stated that when the hip hugger pants became popular they purchased panties with the lower waist to wear with the new pants. Most women expressed great dislike for wearing natural waist panties with lower waist pants. Many women have continued to wear bikini panties even though outer garments now have natural waists.

Internal Search

Women reported that they liked pretty panties and were

heavily influenced by appearance. They usually bought inexpensive panties but occasionally received more expensive ones as gifts from relatives.

External Search

Evaluative criteria used by women when selecting panties were appearance and comfort. As stated before, most women selected panties from those available in stores based on their appearance. Most women preferred panties in a variety of colors or with an applied decoration such as embroidered flowers. A few women preferred white panties.

Women considered comfort important, too, but seldom tried on panties before purchasing. Comfort was judged by visual inspection. Dislike was expressed for elastic that became either too tight or too loose after washing and for seams that scratched. Some women expressed their dissatisfaction with the performance of nylon panties that were dried in a dryer.

Girlfriends or peers were sometimes influential in the selection of panties. The older women indicated that husbands or boyfriends influenced what they purchased.

Purchasing Process

Most women reported that they bought panties in department stores and to a lesser degree in specialty stores. They usually made selections from the tables of panties on display. Some women indicated that they purchased a specific brand such as Sears because they were pleased with previous experiences.

CONCLUSIONS AND FUTURE RESEARCH

A decision process approach revealed substantial insight into the variables that are important in purchase decisions for selected garments. Analysis of focused group interviews, using the Engel-Kollat-Blackwell model, permits a sampling of decision elements in garment decisions and suggests many hypotheses that might be investigated with quantitative or statistical methodologies.

In addition to the specific descriptive elements that might be further investigated, the following general questions are generated from this exploratory research and may be fertile areas for further investigation:

1. The women in this study (who because of selection procedures have above average interest in fashion) appear to be very aware of fashion trends and have well-developed opinions about such trends. Could groups such as these be isolated and with proper statistical controls be used reliably to predict future fashion acceptance and trends?
2. Substantial differences in opinions were expressed by younger and older women regarding some items of clothing. Will the younger women continue to express the same attitudes and beliefs as they grow older (attitude stability) or will their attitudes and beliefs become more consistent with the older women's as they mature? (attitudinal assimilation)
3. Practically every garment was discussed in terms of its comfort by respondents. What is the meaning of comfort? Do women use "comfortable" to describe clothes because it is socially acceptable comment or are physical and performance reasons underlying comfort more important?
4. The current definition of a beautiful person appears to minimize physical features and the clothing worn. Is this definition valid among the entire population?

Will this definition decrease the importance individuals place on clothing? Is this belief in what a beautiful person may be, reflected in the clothes chosen by an individual?

REFERENCES

- Joyce Arbaugh, "Profiling the Textile/Apparel Consumer: A Study of the Usage of Care Label Information," unpublished Ph.D. Dissertation, The Ohio State University, 1974.
- Ruth Ayers et al., Consumer Satisfaction with Men's Shirts and with Women's Slips and Casual Street Dresses, Cornell University Agricultural Experiment Station, Bulletin 984, New York State College of Home Economics, Ithaca, New York, 1963 .
- Barbara Borrell et al., Consumer Satisfaction with Shirts, Slips, Dresses, Cornell University Agricultural Experiment Station Bulletin 985, New York State College of Home Economics, Ithaca, New York, 1974 .
- Marilyn Hunter Braguglia and Mary Lou Rosencranz, "A Comparison of Clothing Attitudes and Ownership of Negro and White Women of Low-Economic Statue," Journal of Consumer Affairs, II(1968), 182-87.
- Yvonne L. Clayton and Lorna R. Sherman, Homemakers' Opinions About Fibers in Selected Household Items: A Nationwide Survey, Marketing Research Report No. 958, Washington: U.S. Dept. of Agriculture, Statistical Reporting Service, 1972 .
- Nancy Burroughs Conklyn, "Consumer Satisfaction with Dress Purchases in a Large Midwest Department Store," unpublished Ph.D. Dissertation, Purdue University, 1972.
- Keith Cox, James B. Higginbotham and John Burton, "Applications of Focus Group Interviews in Marketing," Journal of Marketing, 40(1976), 77-80.
- James F. Engel, David T. Kollat and Roger D. Blackwell, Consumer Behavior, 2nd ed. (Hinsdale, Illinois: Dryden Press, 1973).
- Barbara Good, "A Study of the Textile Product Knowledge of Salespersonnel and Consumer Dissatisfaction with Selected Apparel," M.S. Thesis, The Ohio State University, 1972.
- Martha G. Jenkins and Lois E. Dickey, "Consumer Types Based on Evaluative Criteria Underlying Clothing Decisions," Home Economics Research Journal, 4(1976), 150-162.
- Carolyn Kundel, "Clothing Practices and Preferences of Blue-Collar Workers and Their Families," Home Economics Research Journal, 4(1976), 225-234.
- James Laver, Museum Piece (Boston: Houghton Mifflin Co., 1964).
- Ann L. Orsini, "Search and Transmission of Fabric Performance Information by Consumers and Sales Personnel," unpublished Ph.D. Dissertation, The Ohio State University, 1972.
- Claude R. Martin, Jr., "What Consumer of Fashion Want to Know," Journal of Retailing, 47(1971-1972), 65-94.
- Robert K. Merton, Marjorie Fiske and Patricia L. Kendall, The Focused Interview (Glencoe, Illinois: The Free Press, 1975).
- F. L. Nolan and D. B. Levine, Consumers' Concepts of Fabric, United States Dept. of Agriculture, Marketing Research Report No. 338, Washington: U. S. Government Printing Office, 1959.
- Nellie K. Patson, "Clothing Adequacy of Children Six to Eleven Years Old in Low-Income Families," unpublished Ph.D. Dissertation, The Ohio State University, 1971.
- Mary Ellen Roach, "The Influence of Social Class on Clothing Practices and Orientation at Early Adolescence: A Study of Clothing-Related Behavior of Seventh Grade Girls," unpublished Ph.D. Dissertation, Michigan State University, 1960.
- Mary Ryan, Clothing: A Study in Human Behavior (New York: Holt, Rinehart and Winston, Inc., 1966).
- Jo Ann Schickel, "An Investigation of Selected Factors Related to Clothing and Personal Appearance of Low-Income, Rural Families of Appalachia," unpublished M.S. Thesis, University of Wisconsin, 1970.
- Howard G. Schutz and Beth Ann Phillips, "Consumer Perceptions of Textiles," Home Economics Research Journal, 5(1976), 3-14.
- Florence Skelly, Robert Goldberg and Yvonne L. Clayton, Women's Attitudes Toward Cotton and Other Fibers Used in Wearing Apparel, Marketing Research Report No. 820, Washington: U. S. Dept. of Agriculture, Statistical Reporting Service, 1968.
- Luberta Sims, "Style and Servicability on Clothing Budgets," AATT Technical Review and Register, 1969, 48-9.
- George Byron Sproles, "Consumer Product Attitudes in Market Segmentation Analysis," unpublished Ph.D. Dissertation, Purdue University, 1969.
- James E. Stafford and Ben M. Enis, "The Price-Quality Relationship: An Extension," Journal of Marketing Research, VI(1969), 456-58.
- Linda Ann Stauffer, "Mothers' Use of Consumer Information for Selection and Maintenance of Preschool Children's Clothing," unpublished M.S. Thesis, Virginia Polytechnic and State University, 1972.
- Lynn Barbara Steiniger and Rachel Dardis, "Consumer Textile Complaints," Textile and Chemist and Colorist, III(1971), 33-37.
- Frederick D. Sturdivant, "Low-Income Shoppers in Small Towns: An Exploratory Study of Subcultural Differences in Two Small Towns," Proceedings, Second Annual Conference of the Association for Consumer Research, Edited by David M. Gardner, 1971, 15-55.
- Marjorie Wall, Lois E. Dickey and W. Wayne Talarzyk, "Consumer Satisfaction with Wear and Care Performance of Women's Apparel," Working Paper Series 76-15, College of Administrative Science, The Ohio State University, 1976.
- Margaret Cynthia Warning, "The Implications of Social Class for Clothing Behavior: The Acquisition and Use of Apparel for Girls Seven, Eight and Nine Years of Age in three Social Classes in Des Moines, Iowa," unpublished Ph.D. Dissertation, Michigan State University, 1956.
- Mary Whitlock, Ruth Ayres and Mary Ryan, "Consumer Satisfaction with Women's Blouses, Part I," Field Study in Four Communities in the Northeast, Bulletin 349, Agr. Exp. Sta. Kingston: University of Rhode Island, 1959.

M. J. Wylie, E. M. Crown and M. A. Norris, "Consumer Reaction to Color Performance in Textiles," Home Economics Research Journal , 5(1977), 167-175.

RESEARCH PRIORITIES

Elizabeth Y. Davis, CSRS/U. S. Department of Agriculture
 R. J. Ritchey, Virginia Polytechnic Institute
 & State University
 Francille M. Firebaugh, Ohio State University

Introduction (Elizabeth Y. Davis)

"It is most important to decide what it is you really want to do, and then be realistic enough to recognize how feasible it is to achieve it, and what it takes." This is a direct quotation from the speech by Dr. Walter L. Fishel at the HERAPP Workshop in Washington last spring. HERAPP is the acronym for Home Economics Research Assessment, Planning and Projections. Some of you attended that workshop, or participated in one of the many activities included in the assessment, planning and projections of research.

HERAPP is a product of the ideas and energies of many people. In the summer of 1975 there was a very successful conference where priorities for research in food and agriculture were set through a modified Delphi technique. The Conference was an activity of ARPAC, the Agricultural Research Policy Advisory Committee, made up by joint agreement with the U. S. Department of Agriculture and the National Association of State University & Land-Grant Colleges. In other words, ARPAC represents State University-Federal research relationships. ARPAC is an active committee, and functions as a research planning group in the broadest sense.

When home economics presented their plan for research assessment, planning, and setting of priorities, to ARPAC, it agreed to co-sponsor the project with AAHE (Association of Administrators of Home Economics.)

The relationships among the various organizations represented in the HERAPP project are illustrated below:

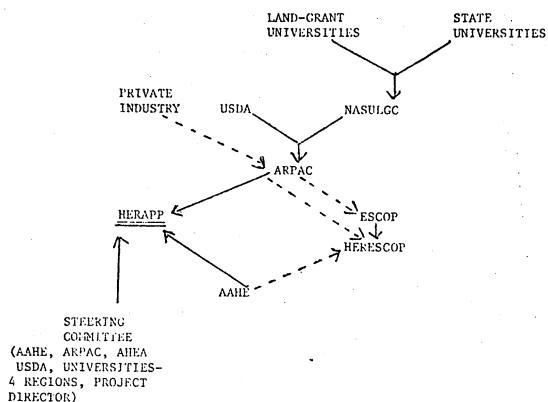


FIGURE 1. Organizations represented in HERAPP project.¹

¹ Abbreviations used are AAHE, Association of Administrators of Home Economics; AHEA, American Home Economics Association; ARPAC, Agricultural Research Policy Advisory Committee; ESCOP, Experiment Station Committee on Organization & Policy; HERAPP, Home Economics Research Assessment, Planning and Projections; HERESCOP, Home Economics Research Sub-committee of Experiment Station Committee on Organization & Policy; NASULGC, National Association of State Universities and Land-Grant Colleges; USDA, United States Department of Agriculture.

The Cooperative State Research Service in the Department of Agriculture developed a cooperative agreement with Virginia Polytechnic Institute and State University to support the project, with Dr. S. J. Ritchey as Project Leader.

Where are we now?

To develop background for setting priorities for home economics research, the administrators of Home Economics Research programs were asked to submit names of scientists and users of research who might prepare statements about the state of the art, or situation statements, in a particular area. There were originally 35 areas included in the initial request. With 35 areas no one can say Home Economics is not multidisciplinary! Requests for preparation of papers were made to five people in each area. The ideas in each area were collated into one paper which was returned to the original authors for review and statements of priority researchable problems; suggestions were incorporated into the working papers prepared for the HERAPP workshop. Nearly 200 people, representing many disciplines and viewpoints attended the Workshop, worked through the papers, and established research priorities.

The most important problems in all areas, as set by the Workshop, were listed in random order and mailed to about 4500 persons including a randomized sampling of members of the American Home Economics Association, those who attended the HERAPP Workshop, and administrators of home economics, to set priorities across areas.

At the same time two other activities were in process. One was to assess the current research actively in home economics and related areas, and to relate these activities. The other activity was a questionnaire sent to all State Experiment Station Directors in order to gain knowledge about the climate for home economics research in experiment stations. A preliminary report on the results of these activities are in the next section of this paper.

The products of the HERAPP exercise will be two published reports. One will include the statements of research and the priority researchable problems and results of the assessment. Projected funding needs in each area will be a part of that publication. The other publication will be a short summary for use in public policy activities.

Many of us who have been working with HERAPP over the past couple of years feel that it is a tool for us to accomplish some of the things in research we want to do -- that we think need to be done -- if we are really dedicated to improving the quality of living for people through research on the family. We also think HERAPP is a tool which is helping us to recognize how feasible it is to achieve those things.

Consumer Research and HERAPP
(S. J. Ritchey)

Consumer research in the broad area of home economics is being evaluated and assessed through a project sponsored by the Cooperative State Research Service, USDA, the Association of Administrators of Home Economics, and AHEA. Preliminary findings can be useful in the context of this workshop and your future endeavors in research.

At the present time, research in all areas of home economics is supported through appropriations to the agricultural experiment stations and the Land-Grant system. Though there is a long history of home economics research in this particular arena, support is quite limited. Support for all home economics research is less than 3 percent of total research through the experiment stations. Moreover, much of the total is expended on food and nutrition, leaving between 1-1.5% for all other areas of home economics. Except for foods and nutrition the number of projects, personnel and financial support are quite limiting.

In almost every measure of support, home economics research lags behind research in the experiment station. Technical and clerical support per scientist, funds per scientist, and funds per projects in home economics lag behind other subject matter areas. This may reflect the poor performance of researchers in home economics and probably reflects the notion that much of the home economics research, particularly the social and economic aspects, do not mesh with the primary mission of the agricultural system. However, administrators perceive home economics as a group of disciplines with potential for service through research. Further, they recognize that research in home economics must be improved by increased funding support, by improving the productivity of researchers, and by improving the available facilities.

In the future, home economics research must receive a higher priority. Our researchers must be more productive, must be more aggressive in seeking funding through the experiment stations and through other granting agencies, both governmental and private. Recommendations from the HERAPP project are that numbers of personnel and funding be increased by factors of two and three, respectively, in the next 10 years. A variety of mechanisms may be important to achieve this goal and the leadership group within home economics must be active in the search for funds, the passage of necessary legislation, the increased productivity of present researchers, and the education of a new generation of outstanding young scholars.

Implementation in Relation to
Interdisciplinary Research Efforts
(Francille M. Firebaugh)

Any planning process is just that -- implementation is not inherent in planning. Research planning can give increased visibility for topics. For researchers, information on priorities can provide impetus for the project development. Administrators will continue to look to researchers to justify areas of high and low priorities with the hope of stimulating an orientation to high priority areas.

Research planning leads to another end -- a delineation of research needs which can be presented to help legislators and others know of our research direction. In working to increase support for our research related to home economics, the need for planned research efforts is

great. Response to legislators concerning research needs for formulating legislation can be stronger as a result of this planning effort.

A research program should include as a primary ingredient the researcher's capabilities and interest, with a clear analysis of importance of the research problems in relation to the needs of society. In 1977, an author proposes: ". . . the problems of society always and necessarily involve many disciplines in their study and solution . . . the best research can be done only in the context of the whole problem" (5:29).

Ten years earlier, another author suggested: "Problem-focused research can be disciplinary: in the sciences of man there is probably not a single discipline that has been able to develop without taking account of the needs of practical action, or being concerned with big social problems" (1:203). I believe contributions can continue from both single disciplines and interdisciplinary efforts.

Research funding has been changing for some period of time. Referring to the shift from grant research to contract research, Rist and others propose:

- 1) ". . . as the Federal government has been able to sharpen its priorities in a number of policy related areas, the scientific community has not kept pace and increasingly has responded in inappropriate ways in terms of proposed research . . . While the demands within agencies for focused research have grown, the academic community has continued to operate in a private, entrepreneurial manner" (4:264).
- 2) ". . . from the Federal view, the assumption is that researchers do not care from where their funds come, so long as they get them and are then left alone" (4:264).
- 3) limited research accountability in grants has influenced the shift to contractual agreements (4:264).

We have seen the RFP (Request for Proposals) come into its element in recent years -- with the statement of the research problem, the target population, and the methodological approach specified. I submit to you that part of that which has come through our ignoring needs of policy makers. We have also seen the rise of contract houses, not associated with university groups who can and do move with greater ease in hiring than universities without the "distraction" and "enrichment" of students and other responsibilities, and with built-in interdisciplinary teams or "purchased" teams. Rist makes a fourth point which seems appropriate for our thinking today:

- 4) "The primacy of the model of the single academic working to locate support for his/her particular project will have to be de-emphasized if academic research is to continue to have Federal support on any scale comparable to that it enjoyed until relatively recently. Cooperation among researchers, the effective use of the supporting services of one or more universities in concert, and the willingness to engage in research that is not oriented toward the development of important contributions to theory or method all appear essential" (4:267).

Political and socio-economic factors affect the extent of scientists participation in policy making in the executive branch of government. "Some have argued that political factors (interparty competition, demands, forms of government, policy-makers' attitudes, etc.) explain more variations in policies" (6:243).

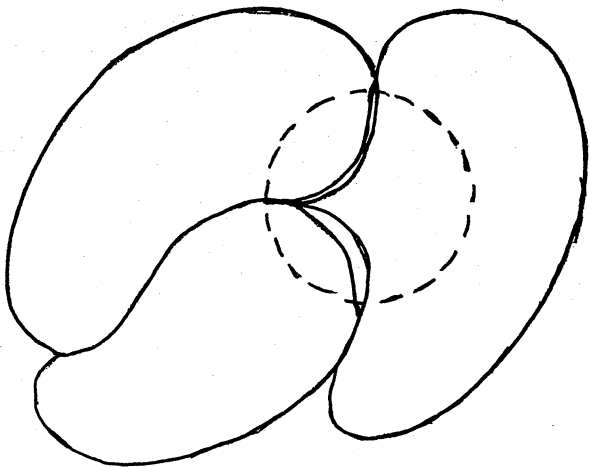
Within marketing, Dyer and Shimp suggest that "policy makers need empirical evidence, and they are receptive to it as long as it captures the full thrust of the issues" (2:67).

We want to be both a part of solutions to problems which are heavily influenced by public policy and we want to be a part of the accumulation and generation of knowledge which depend on continuous contributions to theory.

I would like to differentiate single disciplines, multidisciplines, interdisciplinary and development of new disciplines.

Discipline: Much of the development of knowledge and theory building has come from single disciplines. Perhaps there are individuals who can best function in a single discipline with a broad view of the problem to be researched -- a person with a "cosmopolitan" outlook of the problem but a specific approach which is productive due to the focus.

Multidisciplinary: Boundaries between and among disciplines remain intact, even though cooperation and common attention is focused on a problem or issue. "Mission-oriented work which is so managed as to break down a mission-oriented problem into separate (typically disciplinary) components to be carried out by separate investigators with different skills (possibly) at different sites (5:30). Roy suggests further that multidisciplinary work may take place in widely spread geographic areas: individuals may write reports separately (5:32).



Multidisciplinary

Interdisciplinary: Boundaries are shown here as fading between disciplines as a genuine meshing of ideas takes place. There is a "day-to-day interactive mode of research (or study) where, in order to do the best work, each researcher's work demands the use of ideas, concepts, materials, or instruments from one or more other disciplines. Such research is usually directed to a specified goal or mission" (5:32). Magrabi defined interdisciplinary research as directed toward problems, the solution of which requires application of theories and specialized methods from several disciplines.

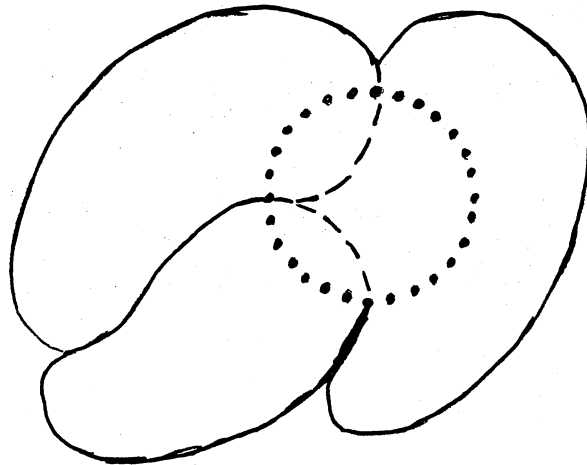
It is fair to ask if it is even possible to determine the research problem as an interdisciplinary group or is it best formulated by one person who seeks help and forms a team. Some have been able to do it, others have not. Luszki, who wrote the classic reference about interdisciplinary team research methods and problems back in 1958 cites four strengths when interdisciplinary groups

formulate research problems:

- 1) resulting research problem should be broader and richer than individuals could construct;
- 2) research problem should be inclusive of important issues which might be overlooked by single researchers;
- 3) research methods from a number of vantage points should be better than those known or experienced by one person; some new concepts may evolve from pooling ideas about the research;
- 4) formulation of problems should assist a researcher in knowing if there is a contribution for him or her to make (3:146).

"When an interdisciplinary team begins work together, difficulty is often experienced in the initial attempt to develop a theoretical framework . . . there may be a tendency to take a set of hunches of hypotheses constituting a loose conglomeration of possible factors and glorify this as theory construction" (3:167). "If the group is successful in developing a theoretical framework to which all participants can subscribe, it changes from a team representing separate disciplines to a cooperative group of individuals with different and overlapping training, concentrating on a common problem" (3:167-168).

In interdisciplinary research, two or more individuals often unite papers or reports together. Roy proposes that program management must be local, "interaction among individual researchers is essential" (5:32).



Interdisciplinary

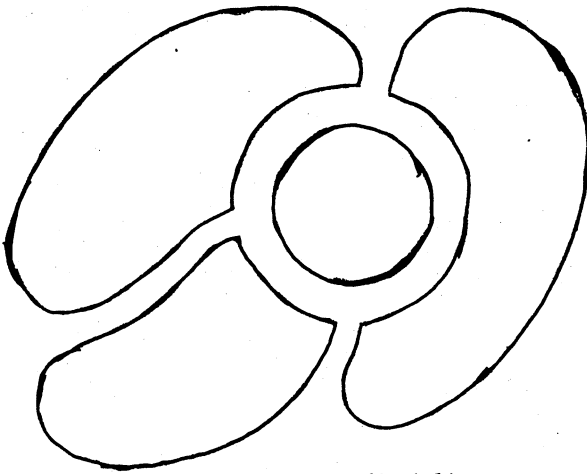
Where does interdisciplinary work lead us? Eventually to new disciplines as noted by Mason earlier today.

New disciplines: "After a generation of concentration of interdisciplinary, concrete issues, new constellations of technical problems will be abstracted out to serve as the foci of new disciplines; and these will then need about thirty years to develop their own specialized theoretical ideas and techniques to the required pitch of excellence, before they in turn are reapplied by a more practically minded generation to the concrete human problems of their time" (7:160).

7. Toulmin, Stephen. From form to function: philosophy and history of science in the 1950's and now. Daedalus, Vol. 106, No. 3, Summer 1977, 143-162.

See also

"Interdisciplinary Research -- An Exploration of Public Policy Issues." Science Policy Research Division, Library of Congress, Serial T, October 30, 1970.



New disciplines

In summary, the Home Economics Research Assessment, Planning, and Projections effort contributes to:

- 1) the stimulus for researching complex issues
- 2) increased visibility of research needs
- 3) foundation for seeking increased funding support
- 4) researcher and administrator information on priorities from several vantage points
- 5) exchange of ideas among researchers -- a starting point for discussion
- 6) recognition of the difficulties of bringing together ideas through brief face-to-face work periods and long distance exchange, with the special challenges of geographical separations and unresolved basic theoretical and vocabulary differences.

Implementation of the research planning will be accomplished by many means, including single, multi- and interdisciplinary programs of research.

References:

1. de Bie, Pierre. Introduction. International Social Science Journal, Vol. 20, No. 2, 1968.
2. Dyer, Robert F. and Terence A. Shimp. Enhancing the role of marketing research in public policy decision making. Journal of Marketing, Vol. 41, No. 1, January 1977, 63-67.
3. Luszki, Margaret Barron. Interdisciplinary team research methods and problems. New York: National Training Laboratories, New York University Press, 1958.
4. Rist, Ray C. Federal funding and social science research: the emergent transformation. Human Organization, Vol. 35, No. 3, Fall 1976, 263-268.
5. Roy, Rustum. Interdisciplinary science on campus -- the elusive dream. Chemical & Engineering News, Vol. 55, No. 35, August 29, 1977, 28-30, 32, 34-36, 38, 40.
6. Schooler, Dean, Jr. Science, Scientists, and Public Policy. New York: The Free Press, 1971.

CONSUMER RESEARCH

Administration and Funding of Interdisciplinary Consumer Research
 Keith N. McFarland, Dean, College of Home Economics
 University of Minnesota

ABSTRACT

Departmental goals and faculty reward systems oriented to department colleagues and activities often impede interdisciplinary efforts in research. Administrative leadership can provide positive reinforcement. Searches for interdisciplinary research funding require institutional, unit administrator and researcher cooperation. The field effort in home economics research is at a relatively underdeveloped stage.

An administrator new to a home economics program quickly becomes sensitive to interdisciplinary endeavor, since the field rationale focuses upon the interrelationship of man and the various aspects of his near environment. Undergraduate courses appear to reach across disciplines in various degree. Interdisciplinary research appears to be more honored in the abstract than in execution. To launch into interdisciplinary endeavor the researcher must form the relationships that permit planning across departmental lines and have the determination that overcomes the added burdens incurred by work outside the department or discipline.

Let's look at some administrative considerations relative to interdisciplinary research. Fred Harvey Harrington describes the department as a "major obstacle to change within American Colleges and Universities."¹ And, indeed, the typical departmental structure poses special problems to those seeking cross-disciplinary outlets. Department goals and individual researcher interest do not always coincide. Staffing needs may be met with choices more appropriate to the disciplinary concerns of the department than to the cross-disciplinary expressions fundamental to a mission-oriented institution. At a time when staff evaluation for tenure, promotion, and salary adjustment purposes is increasingly centered in departmental

¹ Harrington, Fred Harvey, "Structuring of Conventional Departments," from *Academic Departments*, ed. by Dean McHenry, Jossey-Bass Publishers, Inc., San Francisco, CA, 1977.

colleagues, the risks encountered by the cross-disciplinary worker operating "outside the pale," so to speak, are magnified. Seldom are two departments at the same stage of readiness to attack a problem of mutual concern. Other assignments or priorities lay claim to time and resources of those with willingness to work across lines. Those who work in interdisciplinary research must, in a sense, command two or more "languages." In crowded daily routines the time and energy required in setting meetings and in traveling across the campus to work with colleagues of another discipline may be just enough to quell the urge. Outlets for results may be less available for cross-disciplinary material, though this may be changing. The delicate question of who receives the credit in shared research can complicate the scene. When one has limited time for research, efficiency favors its use on matters close at hand. Because cross-disciplinary work may be interpreted as being

peripheral to departmental objectives, the resources of the department may be less available for its promotion.

Some administrative considerations probably necessary to the conduct of interdisciplinary research include the following:

- a) The department must be encouraged to sketch its objectives with sufficient breadth as to include interdisciplinary research as a completely appropriate and, indeed, an expected activity.
- b) The department seminar and project review systems should encourage cooperative interchange with scientists in related disciplines. Joint discussions are useful in legitimizing joint endeavors.
- c) Staff members should be asked to outline research intentions in annual reviews. Discussions attendant to these reviews can lend encouragement and support to new directions in research.
- d) The faculty load must permit the use of reasonable blocks of time for research or for project development and proposal writing. Staff members can be encouraged to get out and make contacts, both professional and funding related. This should be expected behavior. Some travel funds must be available to bring the scientist face-to-face with the potential funding agency. It may be useful to consider redirecting funds presently used in attending society meetings to more focussed visits to funding sources.
- e) The resources of the institution should be utilized to identify possible sources of funding as a preliminary to follow-up visits by the researcher to the agency.
- f) Interest in interdisciplinary research might well become a criterion in faculty recruitment.
- g) Public or private funding agencies need to be made aware of problems of major public interest. Out of this aroused concern should come fiscal response to attractive research proposals. This is a continuing process, never ending.

Funding sources are in the eye of the beholder. Courage, aggressiveness and imagination do on occasion produce strange and wonderful results for the seeker of funds. The researcher in the land grant colleges may have access to Agricultural Experiment Station funds, both federal and state in origin. State funding (legislative specials) is most likely to relate to defined problems and those of interest to legislators who are very close to their constituencies. At the federal level the more comprehensive questions may be treated, and over time one might assume that more funding for consumer research will stem from federal or private than from state sources.

Funding agencies tend to put their money where there is power in program. This works to the disadvantage of new thrusts or of groups just moving into research endeavor. Experienced researchers note the need for

"something to sell" in going to funding sources, this including defined rationale, pilot study results, and the promise of pay-off in ultimate outcomes. Here again the researcher's need for time for planning and for opportunities to meet with funding agencies is evident.

Your workshop planners were interested in the experience of the AHEA Commission of Family Research Act of 1975. The then Senator Walter Mondale had introduced the Family Research Act of 1975 (S2250), with intent to provide funding through the Cooperative State Research Service, USDA for research on the family and to provide assistance to land grant colleges of a kind and duration that would lead to the development of a larger corps of family researchers. It was the commission's role to assist in marshalling such support as we could for this legislation. The commission familiarized itself with the bill and met with the Senator and his staff with respect to language and interpretation. We reviewed our relationships with other interested organizations and agencies, began to structure support groups within each state through the various state Home Economics Associations, and gave thought to potential hearings on this legislation. It quickly became clear that we needed answers to several key questions before testimony could be given with confidence. What is the history of home economics research? What problems are attracting attention? What research would follow funding if the bill were to be enacted? What research of similar nature was being funded by other federal agencies? This information was not at hand.

Land grant college administrators of home economics were asked to report on research programs in progress and to note problems of interest, were funding to be available. Returns were fragmentary, but several interesting findings came to view:

- a) Home economics researchers are few in number. A late inventory of researchers in Agricultural Experiment Stations in 1975 revealed only 153 scientist year equivalents in home economics research endeavor.
- b) The majority of researchers reported investing but limited time in research. The majority of projects reported were of small scale. There was little indication of powerful thrusts in defined areas.
- c) The research projects proposed, were funding to become available, were miscellaneous in nature, and not of a type that would arouse major sponsor support. There were few suggested projects of size and scope. Home economics researchers do, it appears, tend to "think small."
- d) Little was known of the amount of home economics related research funded by other federal agencies.

The fate of the Family Research Act of 1975 (S2250) was sealed by the massive attack launched by various bodies upon the Child and Family Services Act, of which Senator Mondale was also a principal author. Because of the similarity in titles, it was likely that S2250 would, if pushed, attract a parallel attack from conservative groups, hence no hearings were held on S2250 and the bill received no action.

The need for more information on home economics research efforts did lead to the initiation of a study entitled "Home Economics Research Assessment, Planning and Projection" (HERAPP), this funded by the Cooperative State Research Service, USDA, in cooperation with the Association of Administrators of Home Economics and the Agricultural Research Policy Advisory Committee, USDA and NASULGC. The work was headed by S. J. Ritchey and Elizabeth Davis. From their efforts

will come a listing of researchable problems, in priority order. So, too, will the study survey other sources of funding for home economics researchers.

It is quite clear that, in general, home economics researchers have not been aggressive in exploring sources of funding other than that of the Agricultural Experiment Stations. Sources appear to exist. A hasty survey of July and August, 1977 issues of Commerce Business Daily and the Federal Register, reveals 30 requests for proposals and grants in home economics related areas, from USOE, NIAA, NIDA, HEW, NIMH, DOL, PHS, and the U.S. Army. The more complete census being accomplished as a part of the HERAPP study will do much to guide home economists as they look to their research funding future.

The home economics profession has enjoyed some successes in influencing public policy in recent years, yet its history is one of reacting, rather than initiating. This climate is changing.

What lessons did the commission learn from its experience with the Family Research Act?

- a) That a continuing relationship to funding sources needs to be developed and to be maintained over time,
- b) that requests for legislative funding need to rest on a strong information base, be realistic, and promise information useful in dealing with questions held to be important by the funding sources,
- c) that the faculty member needs time to plan, to develop projects and to follow leads to sources,
- d) that the race goes not to the meek or the conservative,
- e) that credibility of the petitioner grows out of action, and that research success breeds success.

Funding of public institutions of higher education increasingly rests upon measures of teaching load. In a future where research will be largely separately funded, the home economics researcher will have to be skillful in seeking research support. Preliminary results from HERAPP suggests that research funding relating to home economics interests is available from a number of Federal Agencies, but that the home economics researcher has to this time concentrated largely upon funding from the CSRS, USDA. A more aggressive inquiry and request program is in order.

EXPERIENCES IN INTERDISCIPLINARY CONSUMER RESEARCH

Sarah L. Manning, Purdue University

Abstract

Experience in two inter-regional interdisciplinary research groups brings out two reasons for doing interdisciplinary research: (1) economic, by pooling funds and costs, and (2) talent-complementing talents and abilities. Frequent problems are (1) communication due to different concepts and word meanings; (2) ego problems with talking prima donnas. Also discussed are publication, administrations, attitude, and concepts/methods problems.

My experience with interdisciplinary research has come from involvement in two regional projects under the Agricultural Experiment Stations of several states. My own field is family economics and management. Other researchers on the committees have been from child development, family studies, housing, and sociology, with some minor nuances of those more interested in management than in economics, or vice versa, those interested in aesthetics rather than sociological or economic factors of housing, those interested in family processes more than the sociology or psychology of the family.

The first committee I was a member of was NC 90, an interdisciplinary, inter-regional research committee to study "Factors Affecting Patterns of Living in Disadvantaged Areas". As this research area was originally conceived to involve more than one discipline, a call went out to all the Agricultural Experiment Station Directors to send a representative if the director and a researcher was interested in being involved. Support for the project came from the several states and also from federal money channeled through the states. One state director was asked to be our administrative advisor. I will quote from the foreword he wrote for the joint publication or basebook we produced:

"It was a rare privilege for me to watch this project develop from the early stages of overcoming the language barrier of diverse disciplines, through the multiple agonies of preparing a common instrument acceptable to all and coordinating the collection and processing of data, to the triumph of the Committee in presenting these data so painstakingly recorded."

Herbert Kramer, Administrative Advisor

Language barriers, multiple agonies, triumph. Herb Kramer's words capture some of our experience. Several members of that committee, believing that we'd been through the language barrier bit already, volunteered to work on a second interdisciplinary committee: NC 128, "Quality of Life as Influenced by Area of Residence." We are currently in the data collection stage, having come through the conceptualization, the development of the instruments and the sampling procedures. My remarks on interdisciplinary research are based on my experience and observations.

First, and also most important, in home economics we are concerned with the family. However you define "family" and its functions, it is a many sided entity, a system composed of subsystems which interact, and a system which interacts with its environment. An interdisciplinary research approach approximates the family better than any one-sided approach can ever do. This is not to

denigrate the one-sided approach; we need researchers going in depth in isolated areas to help develop the building blocks to fit these subsystems together. But the interdisciplinary research offers the best, perhaps the only, way to fit pieces together. As a family economist, I must recognize that, powerful as economics is in shaping family decisions, it is not the only motivating force. My economic analyses may be complete in themselves but they have to take as given tastes, values and social structures. Therefore, as a researcher I welcome the opportunity to work with others who can supply my deficiencies.

Second, on a more practical or personal level, what are our reasons for doing interdisciplinary research?

Economic reasons. We all are constrained by lack of funds. But in NC 90 we were able to obtain thirteen samples totaling 2650 families, with a wealth of detail no one researcher could probably hope to obtain. We decided on central computer processing for our overall analysis. This was done at the University of Missouri for an approximate cost of \$5800. Dr. Edward Metzger estimates that the same processing, done at each of the states, would have cost each of us \$3500 to \$3800 for the computer and other direct costs. We each paid Missouri \$300, so Missouri had to pick up the rest of the tab but even so, Ed estimates that the net cost to Missouri was about \$2200, less than it would have cost them to go it alone. These figures do not cover faculty time and I am sure there was considerable saving of total faculty time over what it would have been without the central processing.

Talent reasons. We each bring to interdisciplinary research our own bag of tricks, our own talents and abilities. Within the committee we find others who complement us by their differences in abilities. The ones who can develop the best conceptual frameworks may not be the ones who have the deepest understanding of computer analysis. The synergistic results are greater than each working alone could have accomplished.

What problems occur in this kind of research? And how can we deal with them?

Communication. As I reflect back to our first formative meeting here in Chicago, I recall that each person acted very protectively for his own research area. There were subtle attempts to put the others down in order to enhance one's own position. In the early meetings we often used the same words and only time showed us that we had different meanings attached to them. How did we overcome this? We read each other's papers and references. We fought with each other, in meetings and over dinner. We liked each other and respected each other so we could bring our differences out into the open. We still haven't arrived at that point where I can say we always resolve our differences, because we haven't. But our meetings have been characterized as "lively". Communication is still a problem but much less so than several years ago.

Ego problems. In commenting on Regional Research in the Southern Cooperative Series, Bulletin 212, October 1976, J.D. Jansma wrote "Faculty members are inherently prima donnas convinced that "their approach" is the preferable one." Not only prima donnas but talkative prima donnas, I believe, for the most part. It takes a firm hand of

chairman to control some of this, and constant checking and rechecking to make sure someone hasn't gone off in his own direction assuming the rest of us are following him. Or not giving a hoot whether we are or not.

Publication problems. Today the pressure is on as never before (and it wasn't negligible before) to publish research. When each state collects data, every researcher wants to work up his own data and get it published. He also would like to pick pieces from the other states to publish. Who has priority when the data came from several or all states involved? In NC 90 we evolved a simple one page form to be sent to each other state when we wanted to ask permission to use any state's data other than our own. This form indicated what parts of the data we wanted, the objectives of the research to be done, the nature of the use or publication and the anticipated completion date. The state wanting the data had 10 days to respond. No response indicated agreement with using the data. Sending to all committee members kept us all informed about what kinds of analyses were being done. If it were something we wanted to do we could either cooperate or shift our emphasis so we didn't have uneconomic duplication. This seemed to solve the problem of individuals publishing by giving them freedom to do so without destroying the unity of the project.

That satisfied the individual researchers, but we still had the obligation to produce a total report from the project. For this purpose we all worked on sections, partially by subcommittees, and then were fortunate to have one committee member who was retiring. She pulled all our parts together into a unified whole. In cooperative research this is a real problem, especially when the committee is large. No one researcher may want to devote so much time to writing the major report when he appears as only one among many authors. Incidentally, one colleague has reported to me that in his university, promotion and tenure committees look at the number of authors and give credit for one third of a refereed article when there are three authors! Not much support for interdisciplinary research evident there.

Administration. This can be either a problem or a help. In general, when interdisciplinary research results from the individual interests of the researchers who see gains to be made either for their field or for themselves, the research will go forward. If it is purely an administrative decision, it seldom works. One administrator several years ago told me that interdisciplinary research did not work because no one took primary responsibility for it; it was always the other fellow's job. This would be less likely to happen if the researchers had initiated it. We have had some people attend our committee meetings who were sent by their Stations. But because there was no time released for them to do research, they dropped out. Or sometimes the administration failed to allocate adequate funds.

Attitude. Any one going into interdisciplinary research would be wise to adopt a cooperative attitude. One must share his ideas, his data, his expertise for the good of the project. This is one of the costs of getting input from others in order that the whole may be greater and that the individual can advance in his field. Compromise is also necessary. Often each of us knows we could go faster or in a deeper direction without all those others we must compromise with. But the synthesizing potential is far greater than going it alone; and without compromise no other researcher worth his salt would be willing to work with you.

Problems of concepts and methods. These differ among various disciplines, and my only advice is to focus on your similarities and your differences until all sides understand. The learning process can sometimes be painful; I hope one of my colleagues from another

discipline has forgiven me for messing up one of his sets of questions in NC 90 when I altered the format to agree with the rest of our questions.

One advantage home economists have in interdisciplinary research is the broad base they already have in their education and orientation. The basic disciplines are fine but we do have this edge, if we are good researchers.

What could improve interdisciplinary research?
Three suggestions:

- 1) Money, of course.
- 2) Time may be more important, however, time to work on the research. Time to interact with one's colleagues on the research team.
- 3) More WATS lines so we could iron out problems faster when we are apart.

INTERDISCIPLINARY RESEARCH NEEDS AND OPPORTUNITIES IN CONSUMER BEHAVIOR AND EDUCATION

J. Barry Mason, University of Alabama
A. William Gustafson, University of Alabama
Mary Harrison, University of Florida
Norleen Ackerman, University of Wisconsin
Roger Swagler, University of Tennessee
Rose Davis, University of Kentucky
Beatrice Petrich, University of Wisconsin
Nancy Miller, University of Wisconsin
E. Thomas Garman, Virginia Polytechnic Institute and State University
Robert Mittelstaedt, University of Nebraska, Lincoln
Terrje Assum, Cornell University

The topics assigned to this ACR/AHEA subgroup for discussion were (1) consumer behavior and satisfaction, (2) family economics, and (3) home economics education, redefined as consumer education. Consumer behavior in the context of the household was the focus of the discussion. The group sought to translate specific aspects of the above three topic areas into a series of researchable topic statements. These statements were then ranked in terms of 1) their need for research and 2) their appropriateness for interdisciplinary research. The background for the three topic areas and the specific research topics agreed upon are presented below.

Consumer Behavior and Satisfaction

Objectives of Research

To strengthen understanding of consumer behavior so as to improve consumer competencies and satisfactions.

Current Situation

One goal of consumer behavior is to maximize satisfaction. Choices by consumers affect the quality of their lives, but many consumers are not adequately informed about the bases of choices or the consequences of decisions. The effects of factors such as governmental regulation, technological developments, and cost and availability of information on consumer satisfaction are not fully known.

Much of the research on consumer satisfaction serves only to determine preferences for brands of products and not whether the product is needed or whether a needed product is available. A better understanding of consumer values and goals and relative degrees of satisfaction would provide a basis for more efficient use of resources. An index of consumer satisfaction, for example, would provide producers and policy makers with new directions for decision making. Policy makers could also be guided in attempts to provide incentives for consumer decisions that conserve nonrenewable resources (for example, through tax credits and deductions) by the results of research on energy use in alternative lifestyles. In an economic system presumably guided by the ideal of maximizing consumer welfare, goals and satisfactions should be the focus of concern in evaluating societal progress.

Researchable Topics

The following topics were identified as those most in need of additional research. (1) The development of a comprehensive conceptualization of the meaning of cost/benefit in the context of consumer protection: (2) the development of measurement tools reflective of the benefits and costs of consumer protection to the consumer: (3) analyze the cost/benefits at the household level of specific consumer protection legislation--i.e.,

fair packaging and labeling, equal credit opportunity and similar legislation.

The most viable interdisciplinary research topics agreed upon were: the same as those developed above as most in need of additional research.

Family Economics

Objective of Research

To analyze the components of economic well-being and evaluate their contribution to the family's overall well-being.

Current Situation

Families are faced with the necessity of reevaluating their expectations with respect to changes in real income and material levels of living. Since World War II families and individuals have experienced a fairly continuous rise in the level of real income. Expectations have developed that this trend will continue bringing possibilities of continued improvement in families' material levels of living. In the past few years, however, the U.S. economy has been faced with both inflation and unemployment which have threatened the maintenance of our currently high material level of living. Also, large numbers of families and individuals in the United States still live in poverty. Whether the size of the economic "pie" continues to increase or remain the same, alternative government programs for income redistribution need to be evaluated.

In addition, the economic significance of family composition and structure is greater today than ever before. The rise in the number of female-headed families, fewer elderly people living with relatives, and fewer couples having children affect the economic well-being of today's family unit.

In the above context, development and use of human capital in the labor market and in the home affect the economy and the individual family unit. Rapid increases in the cost of education, changes in job skills, and increases in female participation in the labor force have repercussions on the economic well-being of families.

Finally, evaluation of economic position, a concept broader than money income, is important, especially because of the increase in the use of credit for consumer purchases and since much of the material level of living is based on possession of durable goods. Economics position of the family, which includes not only the level of money income but debts and savings, and other monetary income are important in determining economic well-being.

Researchable Topics

The following topics were identified as those being most in need of additional research: (1) development of operational definitions of income adequacy and equivalence in order to review appropriateness of poverty thresholds, income supplements, income maintenance and related concepts; (2) studies of economic responses by families to aggregate economic changes including inflation, recession, stagflation, economic dislocations, changing productivity, unemployment and taxes; (3) assessment of the components of levels of well-being of families; (4) an investigation of the impact of public programs and policies on consumption patterns (housing, day care, income supplements, consumer education).

The most viable interdisciplinary research topics were identified as follows: (1) the study of the economic responses by families to aggregate economic changes including inflation, recession, stagflation, economic dislocations, changing productivity, unemployment and taxes; (2) development of operational definitions of income adequacy and equivalence in order to review appropriateness of poverty thresholds, income supplements, and maintenance; (3) study of changes in economic behavior of various groups of families experiencing changes in money income as a result of changes in family composition/circumstances such as multiple job holders and working wives.

Consumer Education

Objectives of Research

To strengthen the system for delivery of knowledge, strengthen the integrative forces within various consumer related disciplines and to improve the education of consumer educators and the public generally.

Current Situation

Consumer education serves individuals and families through professionals who work in such educational areas as:

- 1) higher education: undergraduate and graduate programs
- 2) vocational and technical education: in secondary, post-secondary, adult and paraprofessional programs
- 3) secondary education: in consumer and homemaking, occupational, and general programs
- 4) early childhood and elementary education: in pre-school, kindergarten, primary and middle school programs
- 5) extension, continuing education, and other nonformal programs
- 6) research: in higher education, business, and government
- 7) other, such as in government, business and international programs

Researchable Topics

The following topics were identified as being most in need of additional research: (1) the establishment of systems through which professionals can identify competencies and needs of individuals and families as a base for consumer education programs and devise, validate, and evaluate instruments for use in the systems; (2) determine the characteristics related to the needs of specific learners that can be met by consumer education programs; evaluate models to meet educational needs of these special groups such as displaced homemakers, the elderly, the physically handicapped and institutional (mentally ill) populations; (3) establish systems through which professionals can identify competencies and needs of individuals and families as a base for consumer education programs; devise, validate, and evaluate instru-

ments for use in the system.

The most viable interdisciplinary research topics were identified as: (1) study educational approaches which impede or facilitate change in the consumer behavior of individuals and families in consumer economics related areas (e.g., energy conservation, sex stereotyping, and food habits); (2) determine the characteristics related to the needs of specific learners that can be met by consumer education programs; evaluate models to meet educational needs of special groups such as displaced homemakers, the elderly, the physically handicapped, and institutional (mentally ill) populations; (3) evaluate alternative educational delivery systems in consumer education in relation to such characteristics as purposes, audience, cost, and societal pressure.

Research Topics for ACR/AHEA 1978 Workshop

The following two topics were selected as the subjects for papers to be presented at the ACR/AHEA meeting in June 1978: (1) sources and utilization of information and correlates of satisfaction/dissatisfaction in complaining behavior about food marketing practices at the retail level by the elderly consumer. Authors: Ackerman, Assum, Gustafson, Harrison, and Mason (Hunt and Swagler later joined this group); (2) delineation of professional and subject matter competencies of teachers in consumer education programs. Authors: Davis, Garman, Miller, and Petrick.

Joseph Barry Mason
Reporter

PRIORITY OF TEXTILE AND CLOTHING PROBLEMS
FOR INTERDISCIPLINARY CONSUMER RESEARCH

Gloria Williams, University of Minnesota
Betty Crown, University of Alberta
Marjorie Wall, University of Guelph
Billie G. Murphy, University of North Carolina, Greensboro
Lois Korslund, University of Missouri, Columbia

Six areas were identified in the Home Economics Assessment, Planning and Projections (HERAPP) papers as potentially workable for interdisciplinary research for the Textile and Clothing (T & C) sub group. The areas are (1) clothing--social and psychological aspects, (2) economics of textiles and clothing (3) historic clothing and textiles, (4) design and aesthetics of clothing, (5) special needs for clothing and (6) textiles. Included within each area was a statement of the current situation, a broadly stated research objective, specific research problems of greater and lesser importance and identification of current activity with respect to each problem represented. These papers served as the basis for group discussion.

Suggestions provided by the workshop organizers directed the attention of the Textile and Clothing group to tasks which were primarily evaluative in nature. Initial focus centered on the adequacy of each situation statement--its timeliness, relevance and need for modification. The statement of specific problems for research, the concepts and concepts in relation were the second area of examination. Criteria for assessment included 1) the importance and precedence of research problems within and between the six areas, 2) the significance of the concepts identified, 3) the critical need for research solutions of problems 4) the appropriateness of problems and concepts for interdisciplinary thrusts and 5) the likelihood and necessity of interdisciplinary input into the problem area(s). In addition to suggesting ways in which cooperation among different and interested disciplines could be attained and enhanced, our final task was to identify the significant research problem area(s) or topic(s) within which each group member--separately or together--could develop papers for presentation at the continuing workshop at the 1978 AHEA Conference in New Orleans.

This report addresses itself to the above tasks.

Situation Statements

The subdivisions and titles for T & C areas reflect a traditional approach to categorization of research interest areas in the home economics field. Thus, the possibility of stimulating and attracting interdisciplinary input may be limited. Alert to this likelihood in critical analysis of the situation statements, the following observations and assumptions seem reasonable to make.

1. Current and traditional interests of scientists and educators in the textile and clothing area are reflected in the situation statements. Each, however, has been conceptualized in a manner which lends itself to interdisciplinary input. For example, in AREA I, Clothing--Social and Psychological Aspects,
 - a. attribution theory, function theory and other social interaction theories are implicit in the statement.
 - b. focus is on clothing use which can be conceived as an aspect of general product usage, as an

area of consumer decision-making in post-purchase processes, as a component of non-verbal communication and as a broad area of human behavior (or action) with respect to clothing.

2. Each statement explicitly or implicitly refers to problems connected with theoretical and methodological issues as well pragmatic aspects. It is therefore understood that these issues cannot be separated but must be considered parts of a total perspective for an applied science concerned with consumer research.
3. A stronger practical problem-orientation is apparent in the situation statements of some areas rather than others. For example,
 - a. in AREA VI, Textiles, reference is made to problems connected with (1) inter-and intra-information delivery systems among government, manufacturers/distributors and consumers of textile products, (2) energy resources, use and conservation in the design and recycling of textile structures, (3) the analysis of costs, benefits and consequences of government regulations and standards for producers/distributors and consumers.
 - b. in AREA IV, Design and Aesthetics of Clothing and AREA V, Clothing for Special Needs, attention is directed to problems associated with aesthetic and functional clothing designs for general populations and especially for those populations with special needs--the elderly, children, disabled and obese. Overlap is apparent between these two areas. Marketing problems are evident in production and distribution of clothing for special populations.
 - c. in AREA III, Historic Clothing and Textiles, emphasis is primarily on methodological issues, the need to organize and integrate past knowledge as a basis for describing and explaining the current social-economic-cultural situation and predicting the future along with the importance of preserving these artifacts as a part of our social and cultural heritage.
4. Timeliness and relevance are reflected in certain of the situation statements. For example, emphasis in society on (a) the dignity and worthwhile of individuals and the quality of life experiences (in clothing use, choice and protection) (b) special interest groups--the disabled-elderly, elderly and poor (with respect to special clothing design and quantity needs), women and other minority groups (the stereotypic reflections of appearance and clothing in media presentations and other interacting situation) (c) the proliferation, resistance and acceptance of public policy intervention (in regulating manufacturers and distributors of apparel, home furnishing and other textile products).

Research Problems and Concepts

Taking into consideration the first few research problems identified as more or less important for each subdivision for the T & C group, the following statements seem reasonable:

1. Problem areas as delineated are complex in nature. Contained within each are many concepts and relations between them which are perceived as associative and/or causal; multivariate rather than simplistic relations are to be considered.
2. Some problem areas may be better managed by industry or groups other than those in the disciplines currently represented, e.g., medical and biological teams may be required for some problem areas; some problem areas in textiles are concerned with basic research and development of fiber, yarn, fabric properties.
3. Some problems seem to lean more toward theoretical and methodological issues which seem somewhat removed from providing immediate solutions. Attention to these problems has consequences for building knowledge in the discipline and are basic to solving practical problems.
4. Problem areas through their concepts not only tap different disciplines but different interest and expertise.

In order to tap differences in training, interests and expertise of group members and to provide a means for individual commitment, one approach taken was

1. to develop a matrix which would take into account
 - a. significant and prioritized problem areas on one axis,
 - b. different units of analysis, e.g., consumers, manufacturers, distributors, government, on the other axis and
 - c. interaction between the units and the problem areas
2. to review the lists of researchable problems to see if the concepts and relations exemplified in the problem statements fall more naturally into one cell or another within the matrix.
3. to use the matrix as a basis for subsequent discussion, assessment, commitment and reporting.

The matrix and list of problems for which interdisciplinary research effort is required for problem solutions is attached. (See Appendix).

Individual Topic/Problem Commitment

Reflection on the HERAPP papers, the matrix and activities of the group showed a groping and constant search for a way to capitalize on individual resources and talents as well as manage our tasks of prioritizing research problems and assessing their usefulness for interdisciplinary thrusts. It became apparent in ensuing conversations that each of us was interested in focusing broadly on problems associated with some aspect of the consumer (individuals or families) as the unit of analysis, and his behavior (e.g., use, dressing-undressing) interrelated with aspects of the structure and characteristics of the clothing or flexible product. Consideration would have to be given to conceptual, methodological and pragmatic aspects of the broad problem area or subareas. Three broad subareas were identified for individual pursuit.

1. Problems associated with individual dispositions (i.e., motivations, attitudes and values) with respect to clothing use.
Gloria M. Williams
2. Problems associated with aesthetic and functional clothing design needs with respect to special populations.
Lois Korslund
3. Problems associated with consumer attitudes toward textile product performance and standards.
Betty Crown and Marjorie Wall
(working jointly on product safety)
Billy Murphy

Each of us is aware of the continued and necessary critical assessment of the research problem area and concepts in relation to the specified criteria. In addition, each of us is attuned to the implications of these topics (or areas) for intra- and inter-disciplinary consumer research and possible ways of attaining and enhancing cooperation between disciplines.

Members of the T & C Committee:

- Betty Crown
- Lois Kroslund
- Clark Leavitt
- Billy Murphy
- Fern Rennebohm
- Marjorie Wall
- Gloria Williams

Appendix A

This group identified three main problem areas for which the identification of solutions requires an interdisciplinary research effort (Figure 1). In each area, research is needed to obtain information about consumer needs, values, motives, attitudes and behavior; information about the design, production and distribution of textile and clothing products; and information about government involvement. In addition, there is a need to know more about methods of disseminating the above information among consumers, designers, producers, distributors and government agencies.

Dissemination of Information
(6, 23, 24)

←-----→

PROBLEM AREAS	Consumer needs, values, motives attitudes and behavior	Designers Manufacturers Retailers	Government
Energy/resource use and conservation	*1, 2, 3, 4, 5, 6, 7,	16, 17, 18	2, 22
Product Safety (or more generally, product performance)	8	19	22
Special Needs/Functional Problems	9, 10, 11, 12, 13, 14	16, 20	
	7, 15	21	

Figure 1. Research Problems for Interdisciplinary Consumer Research.

*Numbers in each cell refer to problem statements listed on the following pages.

Problem Statements

Twenty-four sub-problem statements from the HERAPP Draft Working Paper were selected and revised to fit within the framework of Figure 1. The sub-problems thus arrived at are listed below.

1. Examine the effect of fashion on the consumption of clothing and textiles, including prices, quantities of items purchased, and frequency of replacement.
2. Measure the impact of environmental controls on the cost and availability of textiles and clothing; economic implications for clothing from industrial constraints; cost of enforcing and administering the controls; cost and use of care labelling.
3. Examine textiles and clothing consumption from a theoretical view; analyze the clothing and textiles consumption process including cost, discard, and other aspects; relate to demographic and other variables.
4. Analyze consumer demand for clothing and textile products; prices consumers are willing to pay in light of particular product attributes.
5. Develop ways of conserving energy by a use of innovative textile structures in clothing and household textiles and b) modification of care techniques.
6. Evaluation of information delivery systems for transfer of knowledge concerning textile products to the consumer.
7. Determine consumer knowledge and values concerning textile characteristics, and criteria for selection, use and care.
8. Determine consumer knowledge and value systems concerning "trade-offs" involved in textile production, cost, performance (including special properties) and trade resulting from government mandated programs related to consumer and worker safety, environmental protection, and care labelling.
9. Investigate the use of clothing in adapting and adjusting throughout the life cycle.
10. Investigate the interrelationships between textile properties, human anatomy, garment cut, and structure as they affect the comfort, function, and aesthetic quality of garments.
11. Develop measurement methods for assessing garment comfort.
12. Analyze disabilities which create special clothing needs of adults and children and identify problem areas; develop and evaluate solutions to the problem areas identified.
13. Develop a classification system based upon disabling conditions for the synthesis and integration of information about clothing for disabled.
14. Investigate the interrelationships between textile properties and garment design as they affect the physiological and psychological aspects of comfort in varied environments.
15. Develop appropriate techniques for measuring values, attitudes and behavior associated with clothing.
16. Investigate the potential of innovative construction techniques with new fibers and fabric structures for development of new design forms for apparel.
17. Determine the comparative amounts of energy consumed in the production of various types of textile structures.
18. Develop methods for recycling textiles which are profitable economically in order to conserve finite resources.
19. Determine aspects of textiles which affect consumer safety. These include fabric and product flammability as well as the allergenicity, toxicity, mutagenicity, and carcinogenicity of textile additives, dyes, and finishes.
20. Develop a predictive method of shaping and adjusting garments to varied body shapes.
21. Analyze the structural and aesthetic trade-offs made by apparel producers in the interests of controlling production costs and meeting regulations.
22. Determine effects of government policies and regulation relating to textiles on their cost, availability, performance, and selection as well as upon the domestic-international trade balance.
23. Examine the interaction of consumers and retailers in terms of consumer practices, retailer response to these practices, transmission of information and feedback of consumer preferences to producers.
24. Evaluate methods of disseminating to the disabled information on selection, adaptation, and construction of clothing to meet special needs.

INTERDISCIPLINARY RESEARCH GOALS FOR FOOD,
HUMAN NUTRITION, AND HEALTH

Jim McCullough, University of Arizona
Karen Morgan, Michigan State University
Joel Saegert, University of Texas, San Antonio

Problems in the area of Food, Human Nutrition and Health are, by the nature of the field, multidisciplinary, since the disciplines involved in the study of this area are both numerous and diverse. Effective analysis of these problems requires not only multidisciplinary research in which the many disciplines consider a problem from their unique position, but also interdisciplinary research in which both the concepts and viewpoints of the various disciplines are blended to achieve a unique, more complete perspective of the questions in a global setting.

Rather than a single discipline placing emphasis on indepth study of problems drawn from multidisciplinary areas, some segments of the academic community are beginning to recognize the value of an interdisciplinary approach to the study of these problems. It was interest in merging disciplines for the purpose of problem identification that led to the interdisciplinary workshop held by the Association for Consumer Research (ACR) and the American Home Economists Association (AHEA) during the ACR meetings in Chicago.

The purpose of this workshop was to stimulate an interdisciplinary approach to the identification of problem areas from among research issues drawn from several disciplines. One approach to this task is discussed in this paper.

Appropriateness of Problems

Research involving food and nutrition involves numerous disciplines. Table 1 shows some of the disciplines which could make contributions to problem solving in this area of research. Among the 16 fields mentioned in Table 1, only four, i.e., Food Science, Marketing, Psychology, and Economics, were represented by the members of the team preparing this paper. Therefore, many problems, particularly those of a technical nature in the areas of food safety and nutrition, were rejected as inappropriate for treatment by this team.

TABLE 1

Disciplines Related to Food, Human Nutrition, and Health

	<u>Physical Science</u>	
Chemistry		Food Science
	<u>Biological Science</u>	
Biochemistry		Pharmacology
Biophysics		Microbiology
Physiology		Nutrition
	<u>Social Science</u>	
Economics		Anthropology
Sociology		Political Science
	<u>Behavioral Science</u>	
Psychology		Marketing
Management		Dietetics

A three element screening criterion was used to identify those research priorities most suitable for analysis by this particular group. The problem areas most suitable for additional assessment by this team were determined to be:

1. Problems of cognition and perception of food and nutrition
2. Problems of group and individual behavior towards food
3. Problems of attitudes, values and beliefs concerning foods and nutrition.

Problems outside these categories could be better approached by individual scientists in specialized areas or by interdisciplinary teams with greater competence in the more technical areas of food and nutrition research. As a part of an on-going program to establish research priorities in Home Economics, the Home Economics Research Assessment, Planning and Projections (HERAPP) prepared a listing of areas of current research and important problem areas which require additional research. These were screened by the interdisciplinary team using the three element criterion. Since the earlier workers had divided the area of Food, Human Nutrition, and Health into six sub-areas, the problems were examined and classified by the interdisciplinary team under these headings as shown in Table 2. In each area, several problem areas could be appropriately handled, and in most cases, the previously identified major problem areas were found to be feasible concerns. The scheme of classification, however, failed to provide a research frame suitable for systematically identifying appropriate significant research questions for interdisciplinary work; therefore, additional tools were selected.

TABLE 2

Assessment of Problems

<u>Area</u>	<u>HERAPP Problems</u>	<u>Suitable Problems</u>	<u>Most Critical Questions</u>
Food Quality	11	5	Education and Provision of Information
Food Safety	15	2	Education
Food Service	12	10	Energy Conservation Education
Nutrient Requirements	23	4	Biological Requirements Testing
Nutritional Health of Population	9	5	Standard Development, Theory of Behavior
Human Nutrition, Health Education, and Applied Programs	12	12	Theory of Behavior Education

Taxonomy of Projects

Evaluation of the substantive questions underlying the research priorities deemed appropriate for interdisciplinary research resulted in a taxonomy of research areas. This taxonomy is outlined in Table 3.

TABLE 3

A Taxonomy of Problem Areas

- 1) Theoretical Issues Underlying Choice, Habits, and Behavioral Change Relative to Food.
- 2) Analysis of the Vehicles of Change
 - a) information
 - b) intervention
 - c) institutional modification
- 3) Evaluation Criteria

The research proposals from HERAPP deal heavily with methods for change; furthermore, they tend to concentrate on education. Unfortunately, the theoretical bases for these desired changes are not well understood. Consumer information processing and food distribution systems are areas of marketing research which are actively pursued and are areas particularly suitable for interdisciplinary work. The establishment of standards and evaluation criteria must be done in an interdisciplinary fashion if there is to be any real meaning to the results.

The most serious question, however, is the lack of a well-balanced interdisciplinary base from which research may commence. The research question with which this particular team will begin the literature review is "the identification of the factors influencing consumer choices in the food marketplace." This will then serve as a broad base for examination of more specific research questions.

ENERGY MANAGEMENT FOR CITIZENS IN A CONTEMPORARY WORLD

Merlene Lyman, South Dakota State University
Maryann Paynter, University of Kentucky
Emma Auer, Purdue University
Raedene Combs, University of Nebraska
Patricia Tripple, University of Nevada, Reno
Judy Van Name, University of Delaware

Abstract

Work group five, Energy, Shelter and Community, was composed of six persons who explored the assignment and tentatively agreed to compile annotated bibliographies of research related to energy management for citizens in a contemporary world.

Introduction

Work group number five was organized informally during the Friday morning session. Francis Magrabi, co-chairperson of the ACR-AHEA Interdisciplinary Research Workshop, asked conferees to declare a preference from the five research areas she listed. The six persons who chose "Energy, Shelter and the Community" met together to select a discussion leader and recorder as suggested. They held several discussions Friday afternoon and Saturday morning to select and revise research priorities using the preliminary draft of a project report entitled, "Home Economics Research, Assessment, Planning and Projections" (HERAPP).

The research topic areas from the HERAPP project assigned to group five included:

1. Housing and Energy
2. Energy Management in Housing
3. Management and Resource Use
4. Community Development

Initially, participants in this section expressed tentative concerns about:

1. the limited interdisciplinary composition;
2. the assumption that these persons represented the expertise necessary to select and revise the research priorities from the HERAPP project;
3. the scarcity of knowledge concerning the origin or selection of the HERAPP lists (Many participants had access to this information from their institutions but were not aware that this was to be the basis for their deliberation at the workshop); and
4. the lack of direction about how this segment of the workshop was to function during the time before it was expected that they would present a review-of-literature-and-theory paper at the AHEA meeting in New Orleans in June 1978.

Research Priorities

After reading and studying the HERAPP document, suggestions were presented for research priorities. Most of the group favored ranking energy use related to housing as the first priority; others voted for community development or involvement in public policy as a means of influencing consumer energy management.

It was not possible in the time provided to evaluate and re-order all the HERAPP items. However, the following topics were given some precedence over the others.

1. Alternate housing forms
2. Barrier-free housing for elderly and disabled
3. Constraints on achieving home ownership
4. Energy management in the home
5. Extent of discrimination and problems of enforcing laws for ensuring human rights to housing
6. Impact of utilization of home computers
7. Influencing public policy for housing
8. Investments in human capital, new life styles
9. Life cycle cost analyses for housing, furnishings and appliances
10. Other political, economical and social aspects of housing
11. Promotion of democratic decision-making in the home
12. Relationship of public and private energy costs
13. Solar heating and other energy alternatives
14. Spatialization and design of housing

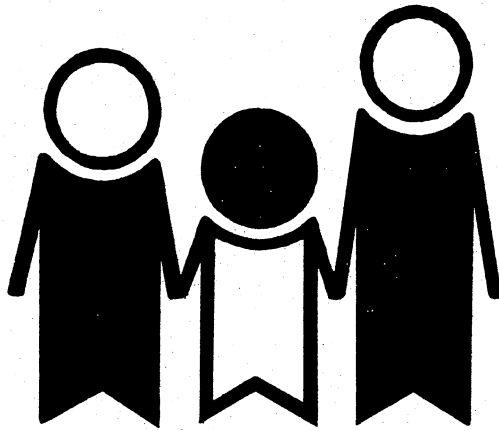
Commitments

Incidences of completed or current projects related to several of the HERAPP researchable problem areas were cited. However, since they were not funded through agriculture funds, it was understood why they were listed. Duplication of these studies did not seem reasonable. Hence, it was decided that each researcher would begin by compiling an annotated bibliography of research using a variety of sources from many disciplines. These are to be sent to Merlene Lyman by February 1, 1978.

Topics selected for the literature search and persons tentatively committed to compiling the annotated bibliographies were:

1. Cost and Availability of Energy for Housing; Adaptation of Technological Advances--
Merlene Lyman, South Dakota State University
2. Contemporary Spatialization and Design of Housing--
Judy Van Name, University of Delaware
3. Relationships Between Public and Private Energy Consumption--
Raedene Combs, University of Nebraska
4. Public Policy Related to Public and Private Energy Management--
Patricia A. Tripple, University of Nevada, Reno
5. Coping Behavior for Managing Family Energy Resources--
Maryann Paynter, University of Kentucky
6. Alteration of Life Styles for Energy Management--
Emma Auer, Purdue University

**Association for Consumer Research
Eighth Annual Conference
O'Hare Hilton
Chicago, Illinois
October 13 - 16, 1977**



H. Keith Hunt, Program Chairman
John Eighmey, Arrangements Chairman
Program Committee:
George Brosseau
Jay Lindquist
Anthony McGann
Earl Morris
William Perreault, Jr.
Robert Pratt, Jr.
Fred Reynolds
William Wilkie
Mary Winter

Program 1977 Annual Conference Association For Consumer Research Thursday, October 13, 1977

Registration opens at 4:00 p. m. on the mezzanine and will remain open until 9:00 p. m.

EARLY-BIRD RECEPTION 6:00-7:30 Cash Bar Terra Firma Room

Evening: Dinner on your own. No evening sessions.

Friday, October 14, 1977

Registration opens at 8:00 a. m. on the mezzanine and will remain open until 6:00 p. m.

Friday Morning Sessions 9:00-10:30

Terra Firma Room

ACR-AHEA INTERDISCIPLINARY RESEARCH WORKSHOP

Chairmen: Frances Magrabi, Consumer and Food Economics, USDA
H. Keith Hunt, Brigham Young University

Note: This is a four session workshop open only by preregistration with Frances Magrabi, Consumer and Food Economics, Federal Building, Hyattsville, Maryland, 20782.

This session covers the problems, satisfactions, and frustrations associated with interdisciplinary research.

Parlor A-B-C

NEW DIRECTIONS IN ATTITUDE RESEARCH

Chairman: Richard Lutz, U.C.L.A.

Discussant: Joel Cohen, University of Florida
Jerry C. Olson, Pennsylvania State University
Philip Dover, Pennsylvania State University

"Attitude Maturation: Changes in Underlying Belief Structures Over Time"

Alan G. Sawyer, Ohio State University
Richard Seminik, University of Utah

"The Temporal Persistence of the Effects of Corrective Advertising"

Barry Schlenker, University of Florida

"Social Identity Theory in Consumer Research"

Richard J. Lutz, UCLA

"A Functional Approach to Consumer Attitude Research"

BREAK 10:30-10:45 on the mezzanine

BREAK 10:30-10:45

- Competitive Paper Session
- # Workshop or Special Topic Session

Chanute/Bleriot Room

MEASURING CONSUMER PREFERENCES *

Chairman: Nancy Thal, University of Colorado
Discussion Leader: Lorraine C. Scarpa, Heublein, Inc.

Thomas J. Stanley, Georgia State University

"Cola Preferences: Disguised Taste vs. Brand Evaluations"

Murphey A. Sewall, State University of New York, Albany

"Nonmetric Multidimensional Scaling of Consumer Preferences for Proposed Product Designs"

Ronald Vaughn, Bradley University
Joseph Pfitlik, Bradley University
Behram Hansotia, Bradley University

"Understanding University Choice: A Multivariate Approach"

BREAK 10:30-10:45

Wright/Farman Room

COMMUNICATION EFFECTS ON CONSUMER BEHAVIOR *

Chairman: Ira Dolich, University of Nebraska-Lincoln

Discussion Leader: Brian Sternthal, Northwestern University

Peter H. Reingen, University of South Carolina

"On the Social Psychology of Giving: Door-In-The-Face and When Even a Penny Helps"

Richard F. Yalch, University of Washington
Margaret C. Dempsey, University of Washington

"Selling a City: An Experimental Study of the Communication Effects of Message Tone"

Linda L. Golden, University of Texas, Austin
Mark I. Alpert, University of Texas, Austin

"The Relative Effectiveness of One-Sided and Two-Sided Communication for Mass Transit Advertising"

BREAK 10:30-10:45

Friday Morning Sessions 10:45-12:15

Terra Firma Room

ACR-AHEA INTERDISCIPLINARY RESEARCH WORKSHOP (continued)

Discussion of research priorities for ACR-AHEA interdisciplinary research.

Groups will be organized to study priorities within special topic areas with reports to be presented Saturday afternoon.

Parlor A-B-C

NEW DIRECTIONS IN ATTITUDE RESEARCH (continued)

Wright/Farman Room

NEW DIRECTIONS IN CONSUMER RESEARCH

Chairman: Ivan Beutler, Iowa State Univ.

Discussion Leader: Kenneth Bernhardt, Georgia State Univ.

Rajendra K. Srivastava, Univ. of Pittsburgh
Allan D. Shocker, Univ. of Pittsburgh
George Day, Univ. of Toronto

"An Exploratory Study of Usage Situational-Influences on the Composition of Competitive Product Markets"

Russell W. Belk, Univ. of Illinois

"An Exploratory Investigation of a Method for Assessing the Effects of Visible Consumption on Impression Formation"

Joseph A. Belonax, Jr., Bowling Green State Univ.
Robert Mittelstaedt, Univ. of Nebraska, Lincoln

"Evoked Set Size as a Function of Number of Choice Criteria and Information Variability"

Chanute/Bleriot Room

CONSUMER RESEARCH IN ADVERTISING AND PROMOTION

Chairman: Jay Lindquist, Western Michigan Univ.

Discussion Leader: Terry Shimp, Univ. of South Carolina

Martin R. Lautman, Assoc. for Res. in Behavior, Inc.
Larry Percy, Gardner Advertising Co.

"Consumer-Oriented versus Advertiser-Oriented Language: Comprehensibility and Salience of the Advertising Message"

Robert Blattberg, Univ. of Chicago
Thomas Buesing, Univ. of Chicago
Peter Peacock, Wake Forest Univ.
Subrata Sen, Univ. of Rochester

"Identifying the Deal Prone Household"

Michael Eigar, State Univ. of New York, Buffalo
Stephen A. Goodwin, State Univ. of New York, Buffalo

"Comparative Advertising: Issues and Problems"

LUNCH 12:30 - 1:50 Galaxy Ballroom

Presidential Address and Special Presentations

Friday Afternoon Sessions 2:00-3:30

Terra Firma Room

INTERDEPENDENT INTERESTS OF CONSUMERS, GOVERNMENT, AND INDUSTRY: THE CASE OF THE TEXTILE INDUSTRY #

Chairman: Christine J. Hager, Economic Research Service, USDA

Rachael Dardis, Univ. of Maryland

"Short and Long Term Cost Consequences of Government Regulations to Consumers"

Hardy Poole, American Textile Manufacturers Institute

"Industry Response to Government Regulations:

George Sproles, Univ. of Houston
Loren Geisfeld, Purdue Univ.

"Consumer Awareness and Satisfaction with Textiles and Apparel"

Jim Sharman, Consumer Product Safety Commission

"The Process of Setting Standards and Some Current Issues in Standards"

BREAK 3:30-3:45

Parlor A-B-C

PERSONAL VALUES AND CONSUMER BEHAVIOR #

Chairman: Donald E. Vinson, Univ. of Southern California

Milton Rokeach, Washington State University

"Personal Values and Human Behavior"

C. Joseph Clawson, Univ. of Southern California

Donald E. Vinson, Univ. of Southern California

"Human Values: A Historical and Interdisciplinary Analysis"

James M. Carman, Univ. of California, Berkeley

"Values and Consumption Patterns: A Closed Loop"

Donald L. Kanter, Univ. of Southern California

"The Europeanizing of America: A Study in Changing Values"

BREAK 3:30-3:45

Wright/Farman Room

RESEARCH ON COGNITIVE STRUCTURE *

Chairman: Steven E. Permut, Yale University

Discussion Leader: William Perreault, University of North Carolina

Jerry Olson, Pennsylvania State University
Daniel Toy, Pennsylvania State University
Philip Dover, Pennsylvania State University

"Mediating Effects of Cognitive Responses to Advertising on Cognitive Structure"

Eric J. Johnson, Carnegie-Mellon University
J. Edward Russo, University of Chicago

"The Use of Response Times to Identify the Organization of Product Information in Memory"

John L. Lastovicka, Temple University
David M. Gardner, University of Illinois

"Low Involvement Versus High Involvement Cognitive Structures"

BREAK 3:30-3:45

Chanute/Bletiot Room

METHODOLOGY — PART I*

Chairman: Kenneth Miller, Univ. of Utah

Discussion Leader: Carol Berning, Procter & Gamble

Morris B. Holbrook, Columbia Univ.
Karl A. Maier, Columbia Univ.

"A Study of the Interface Between Attitude Structure and Information Acquisition Using a Questionnaire-Based Information-Display Sheet"

Robert A. Hansen, Ohio State Univ.
Carol A. Scott, Ohio State Univ.

"Alternative Approaches to Assessing the Quality of Self Report Data"

E. Laird Landon, Jr Univ. of Houston
Sharon K. Banks, Univ. of Oregon

"An Evaluation of Telephone Sampling Designs"

David Sheluga, Purdue Univ.
Jacob Jacoby, Purdue Univ.
Brenda Major, Purdue Univ.

"Whether to Agree-Disagree or Disagree-Agree: The Effects of Anchor Order on Item Response"

BREAK 3:30-3:45

Friday Afternoon Sessions 3:45-5:15

Terra Firma Room

SEX ON TELEVISION #

Chairman: William D. Wells, Needham, Harper & Steers, Advertising

Deborah K. Johnson, Needham, Harper & Steers, Advertising
 Kay Satow, Needham, Harper & Steers, Advertising

"Consumers' Reactions to Sex in TV Commercials"

William Rubens, National Broadcasting Company

"Sex on TV, More or Less"

Susan Hesselbart, Florida State University

"Men and Women on the Tube"

Parlor A-B-C

CONSUMER ORIENTED APPROACHES TO THE STUDY OF TRAVEL BEHAVIOR #

Chairman: Abraham D. Horowitz, General Motors Research Laboratories

Jeffrey S. Milstein, Office of Energy Conservation and Environment, Federal Energy Administration

"Energy Conservation and Travel Behavior"

Alice M. Tybout, Northwestern University
 John R. Hauser, Northwestern University
 Frank S. Koppelman, Northwestern University

"Consumer Oriented Transportation Planning: An Integrated Methodology for Modeling Consumer Perceptions, Preferences, and Behavior"

Jordan J. Louviere, University of Wyoming
 Irwin P. Levin, University of Iowa

"Functional Measurement Analysis of Spatial and Travel Behavior"

John R. Hauser, Northwestern University
 Steve Shugam, University of Rochester

"Interactive Computer Measures of Preference Functions: Applications to Telecommunication and Travel"

Ricardo Dobson, Charles River Assoc., Inc.
 Frederik C. Durbar, Charles River Assoc., Inc.
 N. Scott Cardell, Charles River Assoc., Inc.

"The Hedonic Demand Model: Its Application to Choice of Auto Size"

Mary Lynn Tischer, Urban Planning Division, Federal Highway Administration

"Attitude-Behavior Changes in a 'Before-After' Mode Choice Situation"

Discussants:
 Thomas F. Golob, General Motors Research Laboratories

Mary D. Stearns, Transportation Systems Center, U.S. Dept. of Transportation

Note: This session will run past the 5:15 ending time.

Wright/Farman Room

HALO EFFECTS WITHIN ATTITUDE RESEARCH #

Chairman, Neil E. Beckwith, University of Pennsylvania

Neil E. Beckwith, University of Pennsylvania
 Harold H. Kassarian, U.C.L.A.
 Donald R. Lehman, Columbia University

"Halo Effects in Marketing Research: Review and Prognosis"

Michael J. Ryans, Columbia University

"Halo Effect: Cognitive Structure or Measurement Artifact?"

Joel Huber, Columbia University
 William James, University of Alabama

"A Concept of Halo"

William James, University of Alabama
 Forrest Carter, Georgia Tech. University

"Halo Effects and Location Preferences"

Albert Bemmaro, Purdue University
 Joel Huber, Columbia University

"Specification Error Tests and Halo Effect"

William Moore, Columbia University
 William James, University of Alabama

"A Study of the Amount of Halo and the Perception of Automobiles"

Neil E. Beckwith, University of Pennsylvania
 Victor Kubilius, University of Pennsylvania

"Empirical Evidence of Halo Effects in Store Image Research"

Discussant: William Wilkie, University of Florida

Note: This session will run past the 5:15 ending time.

Chanute/Bleriot Room

METHODOLOGY - PART II *

Chairman: Malcolm White, California State University, Sacramento

Discussion Leader: Simone Clemhout, Cornell University

Sanford L. Grossbart, University of Nebraska, Lincoln
 Robert A. Mittelstaedt, University of Nebraska, Lincoln
 Gene W. Murdock, University of Nebraska, Lincoln

"Nearest Neighbor Analysis: Inferring Behaviors' Processes from Spatial Patterns"

Robert F. Lusch, University of Oklahoma
 Edward F. Stafford, Jr., University of Oklahoma
 Jack J. Kasulis, University of Oklahoma

"Durable Accumulation: An Examination of Priority Patterns"

Richard Mizerski, University of Cincinnati
 Steven Green, University of Cincinnati

"An Investigation into the Causal Links Between Attribution Scheme and Decision-Making"

COCKTAIL HOUR 6:00-7:30 Cash Bar

Terra Firma Room

Dinner on your own. No evening sessions.

Saturday, October 15, 1977

Registration open 8:30 a.m. to 5:00 p.m.

GRADUATE STUDENT BREAKFAST 8:00 a.m. by reservation at the registration desk.

Saturday Morning Sessions 9:00-10:30

Terra Firma Room

HUMAN FACTORS AND CONSUMER RESEARCH

Chairman: J. J. Persensky, National Bureau of Standards

George Whittington, Argonne National Laboratory

"Marketplace Merchandizing of Human Factors"

Lorna Middendorf Consultant to G.M. - Frigidaire

"Fitting the Product to the Person: Human Factors Style"

Tom Cannon, Design Plus
Ronald Hastly, Colorado State Univ.

"Identifying and Defining Consumer Needs Using Human Factors and Market Research Techniques"

Richard G. Snyder, Univ. of Michigan
Clyde L. Owing, Univ. of Michigan
Lawrence L. Schneider, Univ. of Michigan
Herbert M. Reynolds, Univ. of Michigan

"Infant, Child and Teenager Anthropometry for Product Safety Design"

A. M. Ramey Smith, National Bureau of Standards
V. J. Pezoldt, National Bureau of Standards
J. J. Persensky, National Bureau of Standards

"Safety Research — You and Your Lawn Mower"

BREAK 10:30-10:45

Parlor A-B

CONSUMER BEHAVIOR AND HEALTH CARE CHANGE

Chairmen: Thomas S. Robertson, University of Pennsylvania
Lawrence H. Wortzel, Boston University

Thomas S. Robertson, University of Pennsylvania
Terry C. Gleason, University of Pennsylvania
John R. Rossiter, University of Pennsylvania

"Children's Conceptions of Health and Medicine: The Role of Advertising"

M. Venkatesan, University of Oregon

"Nutritional Behavior and Preventive Medicine"

Lawrence H. Wortzel, Boston University
Robert A. Clarke, Boston University

"Consumer Behavior Aspects of Encouraging Non-Smoking"

Irving Wolf, NIMH

"Consumer Behavior and the Prevention of Alcohol Abuse"

BREAK 10:30-10:45

Icarus/Davinci/Montgolfler Room

CONSUMER RESEARCH APPLICATIONS IN PUBLIC POLICY

Chairman: Donald Hempel, University of Connecticut

Discussion Leader: Michael Mazis, Federal Trade Commission

R. Bruce Hutton, University of Denver
Dennis L. McNeill, Arizona State University
William L. Wilkie, University of Florida

"Some Issues in Designing Consumer Information Studies in Public Policy"

Paul N. Bloom, University of Maryland

"Potential Contributions of Consumer Research to Antitrust Decision Making"

Debra L. Scammon, University of Utah

"Effects of Expectancy-Discrepant Information on Consumers"

BREAK 10:30-10:45

Chanute/Bleriot Room

CROSS CULTURAL RESEARCH

Chairman: Donald Hendon, Texas A & I University

Discussion Leader: Roger Blackwell, Ohio State University

Douglas K. Hawes, University of Wyoming
Sigmund Gronmo, Norwegian Fund for Market and Distribution Research
Johan Arndt, Norwegian School of Economics and Business Administration

"Shopping Time and Leisure Time: Some Preliminary Cross-Cultural Comparisons of Time-Budget Expenditures"

J. Michael Munson, University of Santa Clara

"Personal Values: A Cross-Cultural Assessment of Self Values and Values Attributed to a Distant Cultural Stereotype?"

R. Neil Maddox, University of Missouri, St. Louis

Kjell Gronhaug, Norwegian School of Economics and Business Administration
Richard E. Homans, University of Missouri, St. Louis
Frederick E. May, University of Missouri, St. Louis

"Correlates of Information Gathering and Evoked Set Size for New Automobile Purchasers in Norway and the U.S."

BREAK 10:30-10:45

Saturday Morning Sessions 10:45-12:15

Terra Firma Room	Parlor A-B	Icarus/Davinci/Montgolfier Room	Chanute/Bleriot Room
CHILDREN'S INFORMATION PROCESSING AND CONSUMER BEHAVIOR #	ASSESSING AND EXTENDING MULTIATTRIBUTE MODELS*	RESEARCH ON PROGRAMS IN THE PUBLIC SECTOR*	RESEARCH ON LABELING OF CONSUMER PRODUCTS*
Chairmen: Lawrence J. Gianinone, Needham, Harper & Steers, Advertising Paul Zuckerman, Needham, Harper & Steers, Advertising	Chairman: Steve Baumgarten, University of South Florida Discussion Leader: Kenneth W. Kendall, Dalhousie University	Chairman: Betty Diener, Case Western Reserve University Discussion Leader: Richard Reizenstein, University of Tennessee	Chairman: Valerie Valle, University of Pittsburgh Discussion Leader: John Miller, University of Colorado, Colorado Springs
Bobby J. Calder, Northwestern University	James B. Wiley, Wayne State University	Stephen C. Burnett, Indiana University	Philip G. Kuehl, University of Maryland Robert F. Dyer, George Washington University
"Review of Studies on Children's Consumer Socialization"	"Selecting Pareto Optimal Subsets from Multiattribute Alternatives"	"Assessing the Impact of Increased Product Safety on Consumer Utility"	"An Experimental Examination of Deception in Labeling: Consumer Research and Public Policymaking"
Ellen Wartella, Ohio State University	Michel Larocque, Simmons College	Mark I. Alpert, University of Texas, Austin Linda L. Golden, University of Texas, Austin	Roger Best, University of Arizona Jim McCullough, University of Arizona
"Understanding Children's Consumer Socialization"	"Four Methodological Problems in Multiattribute Attitude Models"	"Transportation Attitudes Over Time: A Longitudinal Approach"	"Evaluation of Food Labeling Policies Through Measurement of Consumer Utility"
Marvin Goldberg, McGill University Gerald Gorn, McGill University	Richard A. Werbel, University of Illinois, Chicago Circle	William H. Cunningham, University of Texas, Austin Brondeil Joseph, University of Texas, Austin	F. Robert Dwyer, Jr., University of Arizona
"Television's Role in Promoting Good Eating Habits in Children"	"Measured Attribute Weights Can Make A Difference"	"Energy Conservation: Price Increases and Payback Periods"	"Consumer Processing and Use of Supplementary Drug Label Information"
Discussant: Jim Bettman, UCLA			

LUNCH 12:30-1:50 Galaxy Ballroom

Business meeting will commence during the second half of lunch. Those not interested in attending the annual business meeting should feel free to leave whenever they wish.

Saturday Afternoon Sessions 2:00-3:30

Terra Firma Room

ACR-AHEA INTERDISCIPLINARY RESEARCH WORKSHOP #

Group reports on research priorities for ACR-AHEA interdisciplinary research.

Parlor A-B

RECENT DEVELOPMENTS IN STUDYING CONSUMER INFORMATION PROCESSING #

Chairman: James R. Bettman, UCLA

Jacob Jacoby, Purdue University
Robert Chestnut, Columbia University

"Behavioral Process Technology: Selected Measurement Characteristics"

David Arch, U.C.L.A.
James R. Bettman, U.C.L.A.
Pradeep Kakkar, University of Pennsylvania

"Subjects' Information Processing in Information Display Board Studies"

J. Edward Russo, University of Chicago

"Eye Movements Can Save the World: A Critical Evaluation"

John W. Payne, Duke University
E. K. Easton Ragsdale, University of Chicago

"Verbal Protocols and Direct Observation of Supermarket Shopping Behavior: Some Findings and a Discussion of Methods"

Peter Rip, Stanford University
Peter Wright, Stanford University

"Product Profile Evaluation Tasks in Consumer Judgment Studies"

Meryl P. Gardner, Carnegie-Mellon University
Andrew A. Mitchell, Carnegie-Mellon University
J. Edward Russo, University of Chicago

"Chronometric Analysis: An Introduction and An Application to Low Involvement Perception of Advertisements"

BREAK 3:30-3:45

Icarus/Davinci/Montgolfier Room

RETAILING APPLICATIONS OF CONSUMER RESEARCH #

Chairman: William Locander, University of Houston

Discussion Leader: Ray Marquardt, University of Wyoming

Lawrence J. Ring, University of Virginia
Charles W. King, Purdue University

"A Multiple Discriminant Analysis Approach to the Development of Retail Store Positioning"

James G. Barnes, St. John's University of Newfoundland

"A Hierarchical Model of Source Effect in Retail Newspaper Advertising"

Daniel J. Brown, Oregon State University

"An Examination of Consumer Grocery Store Choice: Considering the Attraction of Size and the Friction of Travel Time"

Chanute/Bleriot Room

CONSUMER SATISFACTION, DISSATISFACTION AND COMPLAINING BEHAVIOR #

Chairman: C. P. Rao, University of Arkansas

Discussion Leader: Earl Morris, Iowa State University

Gerald Zaltman, University of Pittsburgh
Rajendra K. Srivastava, University of Pittsburgh
Rohit Deshpande, University of Pittsburgh

"Perceptions of Unfair Marketing Practices: Consumerism Implications"

Stephen B. Ash, Indiana University

"Rate of Purchase, Importance, and Satisfaction: A Comprehensive Analysis of Durable Goods"

Ralph Day, Indiana University
Muzaffer Bodur, Indiana University

"Consumer Response to Dissatisfaction with Services and Intangibles"

BREAK 3:30-3:45

BREAK 3:30-3:45

Saturday Afternoon 3:45-5:15

Terra Firma Room

ACR-AHEA INTERDISCIPLINARY RESEARCH WORKSHOP # (continued)

Journal editors report their interests in interdisciplinary research.

Uses and funding of ACR-AHEA interdisciplinary research.

Next stop: New Orleans, 1978 AHEA Conference, where the workshop will be continued.

Parlor A-B

RECENT DEVELOPMENTS IN STUDYING CONSUMER INFORMATION PROCESSING # (continued)

Icarus/Davinci, Montgolfier Room

"NEW" MODELS FOR CONSUMER RESEARCH *

Chairman: Gerald Hills, University of Tennessee

Discussion Leader: Ivan Ross, University of Minnesota

Terence A. Oliva, Louisiana State University
Alvin C. Burns, Louisiana State University

"Catastrophe Theory as a Model for Describing Consumer Behavior"

Richard L. Oliver, University of Iowa
Philip K. Berger, University of Kentucky

"Testing Competitive Models of Consumer Decision Making in the Preventative Health Care Marketplace"

Michael J. Ryan, Columbia University

"An Examination of an Alternative Form of the Behavioral Intention Model's Normative Component"

Chanute/Bleriot Room

REPLICATIONS AND EXTENSIONS *

Chairman: George Brooker, State University of New York, Albany

Discussion Leader: Mary Winter, Iowa State University

Philip Kotler, Northwestern University
Lenore Borzak, Northwestern University

"The Market for Personal Growth Services"

Charles M. Schaninger, University of Massachusetts

V. Parker Lessig, University of Kansas

"Sparks and Tucker Revisited: A Reanalysis and Replication"

E. H. Bonfield, Temple University

"Perception of Marital Roles in Decision Processes: Replication and Extension"

COCKTAIL HOUR 6:00-7:15 Cash Bar Terra Firma Room

DINNER 7:30-9:00 Galaxy Ballroom

Sunday, October 16, 1977

Sunday Morning Sessions 9:00-10:30

Terra Firma Room

INTERNATIONAL PAPERS

Chairman: Kenneth P. Uhl, University of Illinois

W. Fred van Raaij, Tilburg University
Kassaye Wandwossen, Tilburg University

"Motivation-Need Theories and
Consumer Behavior"

Solveig Wikstrom, Stockholm University

"Consumer Problems and Consumer Policy
—A Panel Assessment"

Gary Baumgartner, Universite des Sciences
Sociales de Grenoble
Alain Jolibert, Universite des Sciences Sociales
de Grenoble

"The Perception of Foreign Products in France"

Eric Langeard, University of California, Berkeley

"Measurement of Interpersonal Communication
Audiences of Artistic Performance"

Parlor A-B

VISUAL IMAGERY AND ADVERTISING

Chairman: John R. Rossiter

Kathy A. Lutz, U.C.L.A.
Richard J. Lutz, U.C.L.A.

"Illustration and Imagery: Implications for
Advertising"

John R. Rossiter, University of Pennsylvania
Larry H. Percy, Gardner Advertising Co.

"Visual Imaging Ability as a Mediator of
Advertising Response"

Bobby Calder, Northwestern University

"Is There Such A Thing As Imagery?"

Wright/Farman Room

METHODOLOGICAL ASPECTS OF THE ORGANIZATIONAL BUYING PROCESS

Chairman: Robert Spekman, University of
Maryland

Kjell Gronhaug, Norwegian School of
Economics and Business Administration

"Participation in Organizational Buying:
Some Conceptual and Methodological
Problems"

David Wilson, Pennsylvania State University

"Research Approaches to Multi-Participant
Decision Processes"

Arch Woodside, University of South Carolina

"Observations and Organizational Buying
Processes"

Gerald Zaltman, University of Pittsburgh

"Methodological Issues In The Study of
Innovation Adoption by Organization"

Bobby Calder, Northwestern University
Robert Spekman, University of Maryland

"The Collective Level of Analysis: A Crucial
Aspect of Organizational Buying Behavior"

Jerry Wind, University of Pennsylvania

"The Interface Between Organization and
Consumer Buying Behavior"

Icarus/Davinci/Montgolfier Room

NEW DIRECTIONS IN RETAILING RESEARCH: LOCAL MARKET MONITORING

Chairman: Charles King, Purdue University

Douglas Tigert, University of Toronto

"Forecasting Market Shares Through Retail
Image Research"

Larry Ring, University of Virginia

"Defining Market Structures"

George Sproles, University of Houston

"Fashion Preferences and Store Patronage:
A Longitudinal Study"

Dan Emmerheiser, Oklahoma State University
George Sproles, University of Houston

"Segmenting Local Markets for Entertain-
ment Services: The Case of Discoteques"

Charles King, Purdue University

"Strategic Implications of Retailing
Image Research"

BREAK 10:30-10:45

BREAK 10:30-10:45

BREAK 10:30-10:45

BREAK 10:30-10:45

Sunday Morning Sessions 10:45-12:30

Terra Firma Room

CROSS-CULTURAL LIFE STYLE RESEARCH #

Chairman: W. Fred van Raaij, Tilburg Univ.

Johan Arndt, Norwegian School of Economics and Business Administration

Christian Pinson, INSEAD

Manfred Scholer, University of Mannheim

Titles to be announced at session

Discussant: M. Venkatesan, University of Oregon

Parlor A-B

CUE UTILIZATION AND INFERENCE IN CONSUMER BEHAVIOR #

Chairman: Robert E. Burnkrant, University of California, Berkeley

Jerry C. Olson, Pennsylvania State University

"Inferential Belief Formation in the Cue Utilization Process"

Carol A. Scott, Ohio State University

"Self Perception Processes in Consumer Behavior: Interpreting One's Own Experiences"

Bobby Calder, Northwestern University
Alice Tybout, Northwestern University

"New Directions in Attribution Research"

Robert E. Burnkrant, University of California, Berkeley

"Communitization in Product Perception"

Wright/Farman Room

INDIVIDUAL INFLUENCES ON CONSUMER BEHAVIOR *

Chairman: Gerald Alba, University of Oregon

Discussion Leader: Nils-Erik Aaby, Pacific Lutheran University

Elizabeth C. Hirschman, Georgia State University
William O. Adcock, Georgia State University

"An Examination of Innovative Communicators, Opinion Leaders and Innovators for Men's Fashion Apparel"

Robert B. Settle, San Diego State University
Pamela L. Alreck, San Diego State University
John W. Glasheen, San Diego State University

"Individual Time Orientation and Consumer Life Style"

George Belch, U.C.L.A.

"Belief Systems and the Differential Role of the Self-Concept"

James W. Gentry, Kansas State University
Mildred Doering, Kansas State University
Terrence V. O'Brien, Kansas State University

"Masculinity and Femininity Factors in Product Perception and Self Image"

Icarus/Davinci/Montgolfler Room

NEW DIRECTIONS IN RETAILING RESEARCH: LOCAL MARKET MONITORING # (continued)

END OF CONFERENCE

AUTHORS' ADDRESSES

Norleen Ackerman	758	Russell W. Belk	39
University of Wisconsin		School of Commerce and Business Administration	
1270 Linden Drive		University of Illinois, Urbana-Champaign	
Madison, Wisconsin 53706		Urbana, Illinois 61801	
William O. Adcock	308	Joseph A. Belonax, Jr.	48
Department of Marketing		Department of Marketing	
Georgia State University		Bowling Green State University	
Atlanta, Georgia 30303		Bowling Green, Ohio 43403	
Mark I. Alpert	12	Albert C. Bemmaor	477
Department of Marketing	194	Graduate School of Management	
University of Texas, Austin		Purdue University	
Austin, Texas 78712		West Lafayette, Indiana 47907	
Pamela L. Alreck	315	Philip K. Berger	277
School of Business Administration		Department of Business Administration	
San Diego State University		University of Kentucky	
San Diego, California 92182		Lexington, Kentucky 40506	
David C. Arch	555	Roger Best	213
Graduate School of Management		Department of Marketing	
University of California, Los Angeles		University of Arizona	
Los Angeles, California 90024		Tucson, Arizona 85721	
Johan Arndt	151	James R. Bettman	555
Norwegian School of Economics and Bus. Admin.	702	Graduate School of Management	
Helleveien 30	705	University of California, Los Angeles	
N-5000 Bergen, NORWAY		Los Angeles, California 90024	
Stephen J. Arnold	663	Roger D. Blackwell	743
School of Business		College of Administrative Science	
Queen's University		Ohio State University	
Kingston, Ontario, CANADA K7L 3N6		Columbus, Ohio 43210	
Stephen B. Ash	254	Robert Blattberg	57
Marketing Department		Graduate School of Business	
Indiana University		University of Chicago	
Bloomington, Indiana 47401		Chicago, Illinois 60637	
Terrje Assum	758	Paul N. Bloom	138
NYS College of Human Ecology		College of Business and Management	
Cornell University		University of Maryland	
Ithica, New York 14853		College Park, Maryland 20742	
Emma Auer	765	Muzaffer Bodur	263
Department of Consumer Science and Retailing		Marketing Department	
Purdue University		George Mason University	
West Lafayette, Indiana 47907		Fairfax, Virginia 22030	
Sharon K. Banks	103	E. H. Bonfield	300
College of Business Administration		School of Business Administration	
University of Oregon		Temple University	
Eugene, Oregon 97403		Philadelphia, Pennsylvania 19122	
James G. Barnes	235	Lenore Borzak	290
School of Business Administration and Commerce		Graduate School of Management	
Memorial University of Newfoundland		Northwestern University	
St. John's, Newfoundland, CANADA		Evanston, Illinois 60201	
Gary Baumgartner	603	Daniel J. Brown	243
University of Grenoble		School of Business	
Lieu dit "La Marche"		Oregon State University	
69770 Montrouzier, FRANCE		Corvallis, Oregon 97331	
Neil E. Beckwith	465	Thomas Buesing	57
The Wharton School	485	Graduate School of Business	
University of Pennsylvania		University of Chicago	
Philadelphia, Pennsylvania 19104		Chicago, Illinois 60637	
George E. Belch	320	Stephen C. Burnett	186
Graduate School of Management		Marketing Department	
University of California, Los Angeles		Indiana University	
Los Angeles, California 90024		Bloomington, Indiana 47401	

Robert E. Burnkrant	724	Elizabeth Y. Davis	750
School of Business Administration		CSRS	
University of California, Berkeley		U. S. Department of Agriculture	
Berkeley, California 94720		Washington, D. C. 20250	
Alvin C. Burns	273	Rose Davis	758
Department of Marketing		304 Funkhouser	
Louisiana State University		University of Kentucky	
Baton Rouge, Louisiana 70803		Lexington, Kentucky 40506	
Bobby Calder	528	George S. Day	32
Marketing Department	630	University of Toronto	
Northwestern University	653	246 Bloor Street, West	
Evanston, Illinois 60201	721	Toronto, Ontario CANADA M2L 1W5	
Thomas Cannon	494	Ralph Day	263
Design Factors		Department of Marketing	
Fort Collins		Indiana University	
Colorado 80521		Bloomington, Indiana 47401	
N. Scott Cardell	448	Margaret C. Dempsey	5
Charles River Associates		School of Business Administration	
1050 Massachusetts Avenue		University of Washington DJ-10	
Cambridge, Massachusetts 02138		Seattle, Washington 98195	
James Carman	403	Rohit Deshpande	247
Graduate School of Business Administration		Graduate School of Business	
University of California, Berkeley		University of Pittsburgh	
Berkeley, California 94720		Pittsburgh, Pennsylvania 15260	
Forrest S. Carter	474	Lois Dickey	738
College of Industrial Management		Department of Home Economics	
Georgia Institute of Technology		Ohio State University	
Atlanta, Georgia 30332		Columbus, Ohio 43210	
Robert W. Chesnut	546	Nicholas M. Didow	528
Graduate School of Business		Graduate School of Management	
Columbia University		Northwestern University	
New York, New York 10027		Evanston, Illinois 60201	
Roberta Clarke	521	Ricardo Dobson	448
226 Bay State Road		Charles River Associates	
Health Care Management Program		1050 Massachusetts Avenue	
Boston, Massachusetts 02215		Cambridge, Massachusetts 02138	
C. Joseph Clawson	396	Mildred Doering	326
Graduate School of Business Administration		College of Business	
University of Southern California		Kansas State University	
Los Angeles, California 90007		Manhattan, Kansas 66506	
Joel B. Cohen	370	Michael J. Donahue	546
Department of Marketing		Department of Psychological Sciences	
University of Florida		Purdue University	
Gainesville, Florida 32611		West Lafayette, Indiana 47907	
Raedene Combs	765	Philip Dover	78
Home Economics 132		School of Business Administration	333
University of Nebraska, Lincoln		Dartmouth College	
Lincoln, Nebraska 68503		Hanover, New Hampshire 03755	
M. Crousillat	606	Frederick C. Dunbar	448
Universite d'Aix-Marseille		Charles River Associates	
Place Victor Hugo		1050 Massachusetts Avenue	
13331 Marseille Cedex 3 FRANCE		Cambridge, Massachusetts 02138	
Betty Crown	760	F. Robert Dwyer	220
Faculty of Home Economics		Department of Marketing	
University of Alberta		University of Arizona	
Edmonton, Alberta, CANADA T6G 2E2		Tucson, Arizona 85721	
William H. Cunningham	201	Robert F. Dyer	206
Graduate School of Business		Department of Business Administration	
University of Texas, Austin		George Washington University	
Austin, Texas 78712		Washington, D. C. 20006	
Rachael Dardis	377	Daniel A. Emenheiser	682
3103 Turner Lab		School of Hotel and Restaurant Administration	
University of Maryland		Oklahoma State University	
College Park, Maryland 20740		Stillwater, Oklahoma 74074	

Michael Etgar	63	Stephen Green	126
School of Management		Department of Marketing	
State University of New York, Buffalo		University of Cincinnati	
Buffalo, New York 14214		Cincinnati, Ohio 45221	
Karen Kaigler-Evans	738	Kjell Grønhaug	167
Department of Home Economics		Norwegian School of Economics and Bus. Admin.	635
University of Texas, Austin		Helleveien 30	
Austin, Texas 78712		N-5000 Bergen, NORWAY	
Francille M. Firebaugh	750	Sigmund Grønmo	151
School of Home Economics		Norwegian Fund for Market and Distribution Res.	
Ohio State University		Karl Johans GT 4	
Columbus, Ohio 43210		Oslo, NORWAY	
David M. Gardner	87	Sanford L. Grossbart	114
Department of Business Administration		Department of Marketing	
University of Illinois, Urbana-Champaign		University of Nebraska, Lincoln	
Urbana, Illinois 61801		Lincoln, Nebraska 68588	
Meryl P. Gardner	581	A. William Gustafson	758
Graduate School of Industrial Administration		Department of Consumer Sciences, Box 1488	
Carnegie-Mellon University		University of Alabama	
Pittsburgh, Pennsylvania 15213		University, Alabama 35486	
E. Thomas Garman	758	Robert A. Hansen	99
Wallace 226		College of Administrative Science	
Virginia Polytechnic Institute and State Univ.		Ohio State University	
Blacksburg, Virginia 24060		Columbus, Ohio 43210	
Loren V. Geistfeld	383	Behram Hansotia	26
Department of Consumer Science		Marketing Department	
Purdue University		Bradley University	
West Lafayette, Indiana 47907		Peoria, Illinois 61625	
James Gentry	326	Mary Harrison	758
College of Business		3008 McCarty Hall	
Kansas State University		University of Florida	
Manhattan, Kansas 66506		Gainesville, Florida 23611	
Wendy Gibson	540	Ronald W. Hasty	494
Stanford University		College of Business	
Stanford		Colorado State University	
California 94305		Fort Collins, Colorado 80521	
John W. Glasheen	315	John Hauser	426
School of Business Administration		Graduate School of Management	440
San Diego State University		Northwestern University	
San Diego, California 92182		Evanston, Illinois 60201	
Terry C. Gleason	515	Douglass K. Hawes	151
The Wharton School		Department of Business Administration	
University of Pennsylvania		University of Wyoming	
Philadelphia, Pennsylvania 19104		Laramie, Wyoming 82071	
Helen Goetz	734	Susan Hesselbart	419
Box 6195		Florida State University	
University of Alabama		Tampa	
University, Alabama 35486		Florida 33620	
Marvin E. Goldberg	540	David Hillier	380
Faculty of Management, McGill University		American Textile Manufacturers Institute	
1001 Sherbrooke St., West		1150 17th Street, N.W.	
Montreal, Quebec, CANADA H3A 1G5		Washington, D. C. 20207	
Linda Golden	12	Jo Ann Schickel Hilliker	743
Department of Marketing	194	Department of Textiles and Clothing	
University of Texas, Austin		Ohio State University	
Austin, Texas 78712		Columbus, Ohio 43210	
Stephen A. Goodwin	63	Elizabeth C. Hirschman	308
School of Management		Graduate School of Business	
State University of New York, Buffalo		University of Pittsburgh	
Buffalo, New York 14214		Pittsburgh, Pennsylvania 15260	
Gerald J. Gorn	540	Morris B. Holbrook	93
Faculty of Management, McGill University		Graduate School of Business	
1001 Sherbrooke St., West		Columbia University	
Montreal, Quebec, CANADA H3A 1G5		New York, New York 10027	

Oddvar Holmer	702	Charles W. King	227
Norwegian School of Economics and Bus. Admin. Helleveien 30 N-5000 Bergen, NORWAY		Graduate School of Management Purdue University West Lafayette, Indiana 47906	688
Richard E. Homans	167	Frank S. Koppleman	426
School of Business Administration University of Missouri, St. Louis St. Louis, Missouri 63121		Graduate School of Management Northwestern University Evanston, Illinois 60201	
Wayne D. Hoyer	546	Lois Korslund	760
Department of Psychological Sciences Purdue University West Lafayette, Indiana 47907		Clothing and Textiles, College of Home Economics University of Missouri, Columbia Columbia, Missouri 65201	
Joel Huber	468	Philip Kotler	290
Graduate School of Business Columbia University New York, New York 10027	477	Graduate School of Management Northwestern University Evanston, Illinois 60201	
H. Keith Hunt	iv	Victor Kubilius	485
Department of Business Management Brigham Young University Provo, Utah 84602	233	The Wharton School University of Pennsylvania Philadelphia, Pennsylvania 19104	
R. Bruce Hutton	131	Philip G. Kuehl	206
College of Business Administration University of Denver Denver, Colorado 80208		College of Business and Management University of Maryland College Park, Maryland 20742	
Jacob Jacoby	109	E. Laird Landon, Jr.	103
Department of Psychological Sciences Purdue University West Lafayette, Indiana 47907	546	Department of Marketing University of Houston Houston, Texas 77004	
William James	468	Eric Langeard	606
Department of Marketing University of Alabama University, Alabama 35486	474 481	School of Business Administration University of California, Berkeley Berkeley, California 94720	
Deborah K. Johnson	411	Michel Laroche	175
Needham, Harper & Steers Advertising, Inc. 401 North Michigan Avenue Chicago, Illinois 60611		Simmons College Boston Massachusetts 02146	
Eric J. Johnson	79	John Lastovicka	87
Graduate School of Industrial Administration Carnegie-Mellon University Pittsburgh, Pennsylvania 15213		Department of Marketing Temple University Philadelphia, Pennsylvania 19122	
Alain Jolibert	603	Martin Lautman	52
Cidex 633 - le Brinchet St. Nazaire-les-Eymes 38330 St. Ismier, FRANCE		Associates for Research in Behavior 3401 Market Street Philadelphia, Pennsylvania 19104	
Brondel Joseph	201	Clark Leavitt	738
Department of Marketing University of Illinois, Chicago Circle Chicago, Illinois 60610		College of Administrative Science Ohio State University Columbus, Ohio 43210	
Pradeep Kakkar	555	Donald R. Lehman	465
The Wharton School University of Pennsylvania Philadelphia, Pennsylvania 19104		507 Uris Hall Columbia University New York, New York 10027	
Donald Kanter	408	V. Parker Lessig	295
Graduate School of Business Administration University of Southern California Los Angeles, California 90007		School of Business University of Kansas Lawrence, Kansas 66045	
Harold H. Kassarian	xiii	Irwin P. Levin	435
Graduate School of Management University of California, Los Angeles Los Angeles, California 90024	465	Department of Geography University of Iowa Iowa City, Iowa 52242	
Jack J. Kasulis	119	Jordan J. Louviere	435
College of Business Administration University of Oklahoma Norman, Oklahoma 73019		Department of Geography University of Wyoming Laramie, Wyoming 82071	

Robert F. Lusch	119	Jeffery S. Milstein	422
College of Business Administration		Office of Energy Conservation and Environment	
University of Oklahoma		Federal Energy Administration	
Norman, Oklahoma 73019		Washington, D. C. 20461	
Kathy Lutz	611	Andrew A. Mitchell	581
20635 Eagle Pass Drive		Graduate School of Industrial Administration	
Malibu		Carnegie-Mellon University	
California 90265		Pittsburgh, Pennsylvania 15213	
Richard J. Lutz	360	Robert A. Mittlestaedt	48
Graduate School of Management	611	College of Business Administration	114
University of California, Los Angeles		University of Nebraska, Lincoln	758
Los Angeles, California 90024		Lincoln, Nebraska 68588	
Merlene Lyman	765	Richard Mizerski	126
Textiles, Clothing and Interior Design		Department of Marketing	
South Dakota State University		University of Cincinnati	
Brookings, South Dakota 57006		Cincinnati, Ohio 45221	
Sylvia Ma	663	William Moore	481
Faculty of Management Studies		523 Uris Hall	
University of Toronto		Columbia University	
Toronto, Ontario CANADA M5S 1W5		New York, New York 10027	
R. Neil Maddox	167	Karen Morgan	763
School of Business Administration		237 Food Science Building	
University of Missouri, St. Louis		Michigan State University	
St. Louis, Missouri 63121		East Lansing, Michigan 48824	
Karl A. Maier	93	J. Michael Munson	160
Ketchum, MacLeod & Grove, Inc.		Department of Marketing	
90 Park Avenue		University of Santa Clara	
New York, New York 10016		Santa Clara, California 95053	
Sarah Manning	756	Gene W. Murdock	114
School of Consumer and Family Sciences		College of Business Administration	
Purdue University		University of Nebraska, Lincoln	
West Lafayette, Indiana 47907		Lincoln, Nebraska 68588	
Brenda Major	109	Billie G. Murphy	760
Department of Psychological Sciences		Clothing and Textiles Department	
Purdue University		University of North Carolina, Greensboro	
West Lafayette, Indiana 47907		Greensboro, North Carolina 27411	
J. Barry Mason	734	Terrence V. O'Brien	326
Box J	758	College of Business	
University of Alabama		Kansas State University	
University, Alabama 35486		Manhattan, Kansas 66506	
Frederick E. May	167	Terence A. Oliva	273
School of Business Administration		Department of Management	
University of Missouri, St. Louis		Louisiana State University	
St. Louis, Missouri 63121		Baton Rouge, Louisiana 70803	
Jim McCullough	213	Richard L. Oliver	277
Department of Marketing	763	Department of Business Administration	
University of Arizona		University of Iowa	
Tucson, Arizona 85721		Iowa City, Iowa 52242	
Keith N. McFarland	754	Jerry C. Olson	72
College of Home Economics		707-A Business Administration Building	333
University of Minnesota		Pennsylvania State University	706
St. Paul, Minnesota 55416		University Park, Pennsylvania 16802	
Shelby H. McIntyre	160	Clyde L. Owings	499
Department of Marketing		Biomedical Department	
University of Santa Clara		University of Michigan	
Santa Clara, California 95053		Ann Arbor, Michigan 48105	
Dennis L. McNeill	131	John W. Payne	571
Division of National Advertising		Graduate School of Business	
Federal Trade Commission		University of Chicago	
Washington, D. C. 20580		Chicago, Illinois 60637	
Nancy Miller	758	Maryann Paynter	765
University of Wisconsin		319 Funkhouser Building	
1270 Linden Drive		University of Kentucky	
Madison, Wisconsin 53706		Lexington, Kentucky 40506	

Peter Peacock	57	Michael J. Ryan	283
Graduate School of Management		Graduate School of Business	
Wake Forest University		Columbia University	
Winston-Salem, North Carolina 27109		New York, New York 10027	
Larry Percy	52	Joel Saegert	730
24 Forest Oak Court	621	Division of Management and Marketing	763
Fenton		University of Texas, San Antonio	
Missouri 63026		San Antonio, Texas 78285	
Jay J. Persensky	508	Merry Mayne Saegert	730
A353 Building 220		Health Science Center	
National Bureau of Standards		University of Texas, San Antonio	
Washington, D. C. 20234		San Antonio, Texas 78285	
Val J. Pezoldt	508	Kay Satow	411
A353 Metrology Building		Needham, Harper & Steers Advertising, Inc.	
National Bureau of Standards		401 North Michigan Avenue	
Washington, D. C. 20234		Chicago, Illinois 60611	
Joseph Pitlik	26	Alan G. Sawyer	343
Marketing Department		College of Administrative Science	
Bradley University		Ohio State University	
Peoria, Illinois 61625		Columbus, Ohio 43210	
E. K. Easton Ragsdale	571	Debra Scammon	145
Graduate School of Business		Department of Marketing	
University of Chicago		University of Utah	
Chicago, Illinois 60637		Salt Lake City, Utah 84112	
Ann M. RameySmith	508	Charles M. Schaninger	295
A353 Metrology Building		College of Business Administration	
National Bureau of Standards		University of Massachusetts	
Washington, D. C. 20234		Amherst, Massachusetts 01002	
Peter H. Reingen	1	Barry R. Schlenker	352
College of Business Administration		Department of Psychology	
University of South Carolina		University of Florida	
Columbia, South Carolina 29208		Gainesville, Florida 32611	
Lawrence J. Ring	227	Lawrence W. Schneider	499
Graduate School of Business Administration	668	Biomedical Department	
University of Virginia		University of Michigan	
Charlottesville, Virginia 22906		Ann Arbor, Michigan 48105	
Peter Rip	578	Carol A. Scott	99
Graduate School of Business		College of Administrative Science	714
Stanford University		Ohio State University	
Stanford, California 94303		Columbus, Ohio 43210	
R. J. Ritchie	750	Richard Semenik	343
College of Home Economics		College of Business	
Virginia Polytechnic Institute and State Univ.		University of Utah	
Blacksburg, Virginia 24060		Salt Lake City, Utah 84112	
Thomas S. Robertson	515	Subrata Sen	57
The Wharton School	525	Graduate School of Management	
University of Pennsylvania		University of Rochester	
Philadelphia, Pennsylvania 19104		Rochester, New York 14627	
Deborah L. Roeder	528	Robert B. Settle	315
Graduate School of Management		School of Business Administration	
Northwestern University		San Diego State University	
Evanston, Illinois 60201		San Diego, California 92182	
John R. Rossiter	515	Murphy A. Sewall	22
The Wharton School	621	School of Business	
University of Pennsylvania		State University of New York, Albany	
Philadelphia, Pennsylvania 19174		Albany, New York 12222	
William S. Rubens	115	L. J. Sharman	392
National Broadcasting Company		Consumer Product Safety Commission	
160 Chestnut Drive		Washington, D. C. 20207	
Roslyn, New York 11576			
J. Edward Russo	79	Carl G. Shea	456
Graduate School of Business	561	HHP 22, U.S. Department of Transportation	
University of Chicago	581	400 Seventh Street, S.W.	
Chicago, Illinois 60637		Washington, D. C. 20590	

David Sheluga	109	Alice M. Tybout	426
Department of Psychological Sciences	546	Graduate School of Management	721
Purdue University		Northwestern University	
West Lafayette, Indiana 47907		Evanston, Illinois 60201	
Allan D. Shocker	32	Judy Van Name	765
Graduate School of Business		College of Home Economics	
University of Pittsburgh		University of Delaware	
Pittsburgh, Pennsylvania 15260		Newark, Delaware 19711	
Steven M. Shugan	440	W. Fred van Raaij	590
Graduate School of Management		Economic Psychology, G 703	693
University of Rochester		Tilburg University	
Rochester, New York 14627		Tilburg, THE NETHERLANDS	
Richard G. Snyder	499	Ronald Vaughn	26
Highway Safety Research Institute		Marketing Department	
Huron Parkway and Baster Road		Bradley University	
Ann Arbor, Michigan 48105		Peoria, Illinois 61625	
Robert E. Spekman	653	M. Venkatesan	518
College of Business and Management		College of Business Administration	704
University of Maryland		University of Oregon	
College Park, Maryland 20742		Eugene, Oregon 97403	
George B. Sproles	383	Donald E. Vinson	396
Department of Human Development and Cons. Sci.	675	Graduate School of Business Administration	
University of Houston	682	University of Southern California	
Houston, Texas 77004		Los Angeles, California 90007	
Rahendra K. Srivastava	32	Daniel B. Wackman	535
Graduate School of Business	247	University of Minnesota	
University of Pittsburgh		Minneapolis	
Pittsburgh, Pennsylvania 15260		Minnesota 55455	
Edward F. Stafford, Jr.	119	Marjorie Wall	760
College of Business Administration		Department of Consumer Studies	
University of Oklahoma		University of Guelph	
Norman, Oklahoma 73019		Guelph, Ontario, CANADA N1G 2W1	
Thomas J. Stanley	19	Kassaye Wandwossen	590
Department of Marketing		Economic Psychology G 703	
Georgia State University		Tilburg University	
Atlanta, Georgia 30303		Tilburg, THE NETHERLANDS	
Brian Sternthal	721	Scott Ward	535
Graduate School of Management		Harvard Business School	
Northwestern University		Soldiers Field	
Evanston, Illinois 60201		Boston, Massachusetts 02163	
Roger Swagler	758	Ellen Wartella	535
1218 White Avenue		Ohio State University	
University of Tennessee		Columbus	
Knoxville, Tennessee 37916		Ohio 43210	
James L. Taylor	643	R. Weisz	606
Department of Marketing		Universite d'Aix-Marseille	
Texas A & M University		Place Victor Hugo	
College Station, Texas 77841		13331 Marseille Cedex 3 FRANCE	
Douglas J. Tigert	663	Richard A. Werbel	181
Faculty of Management Studies		University of Illinois, Chicago Circle	
University of Toronto		403 North Grove Avenue	
Toronto, Ontario, CANADA M5S 1W5		Oak Park, Illinois 60302	
Mary Lynn Tischer	456	Solveig Wikstrom	596
HHP-22, U.S. Department of Transportation		Department of Business Administration	
400 Seventh Street, S.W.		Stockholm University	
Washington, D. C. 20590		S-10691 Stockholm, SWEDEN	
Daniel Toy	72	James B. Wiley	171
Graduate School of Business		300 Prentis Hall	
University of Washington		Wayne State University	
Seattle, Washington 98195		Detroit, Michigan 48201	
Patricia Tripple	765	William L. Wilkie	131
120 Redstone Drive		Marketing Department	
Reno, Nevada 89512		University of Florida	
		Gainesville, Florida 32611	

Gloria Williams 760
College of Home Economics
University of Minnesota
St. Paul, Minnesota 55416

David T. Wilson 639
707G College of Business Building
Pennsylvania State University
University Park, Pennsylvania 16802

Yoram Wind 657
The Wharton School
University of Pennsylvania
Philadelphia, Pennsylvania 19174

Arch G. Woodside 643
College of Business Administration
University of South Carolina
Columbia, South Carolina 29208

Lawrence H. Wortzel 521
39 Stafford Road
Newton Centre
Massachusetts 02159

Peter Wright 578
Graduate School of Business
Stanford University
Stanford, California 94303

Richard F. Yalch 5
School of Business Administration
University of Washington DJ-10
Seattle, Washington 98195

Eleanor Young 730
Health Science Center
University of Texas, San Antonio
San Antonio, Texas 78285

Gerald Zaltman 247
Graduate School of Business
University of Pittsburgh
Pittsburgh, Pennsylvania 15260

REVIEWERS OF COMPETITIVE PAPERS

Nils-Erik Aaby, Pacific Lutheran University
Gerald Albaum, University of Oregon
David Barnaby, University of Tennessee
Penny Baron, University of Iowa
Steven Baumgarten, University of South Florida
Neil Beckwith, University of Pennsylvania
Russell W. Belk, University of Illinois
Carol Berning, Procter and Gamble Co.
James R. Bettman, University of California, Los Angeles
Ivan Beutler, Iowa State University
Paul N. Bloom, University of Maryland
George Brooker, State University of New York, Albany
Daniel Brown, Oregon State University
Robert E. Burnkrant, University of California at Berkeley
Bob Cecil, Leo Burnett
Simone Clemhout, Cornell University
Ralph Day, Indiana University
Barbara Deering, United Dairy Industry Association
John Gifford, Miami University
Linda L. Golden, University of Texas at Austin
Douglass K. Hawes, University of Wyoming
Donald Hempel, University of Connecticut
Donald Hendon, Texas A & I University
Gerald Hill, University of Tennessee
Harold H. Kassarian, University of California, Los Angeles
Jack J. Kasulis, University of Oklahoma
William Katovsky, Postal Rate Commission
C. L. Kendall, University of North Carolina
Philip G. Kuehl, University of Maryland
E. Laird Landon, Jr., University of Houston
William Locander, University of Houston
Richard Lutz, University of California, Los Angeles
R. Neil Maddox, University of Missouri, St. Louis
J. Barry Mason, University of Alabama
Ray Marquardt, University of Wyoming
Michael Mazis, Federal Trade Commission
John Miller, University of Colorado, Colorado Springs
Ken Miller, University of Utah
Ken Monroe, Virginia Polytechnic Institute
Richard L. Oliver, University of Kentucky
Richard Olshavsky, Indiana University
Jerry C. Olson, Pennsylvania State University
Larry Percy, Gardner Advertising Com.
Steven Permut, Yale University
Jay J. Persensky, National Bureau of Standards
Richard Reizenstein, University of Tennessee
Frederich Russ, University of North Carolina
Michael J. Ryan, Columbia University
Alan Sawyer, Ohio State University
Debra L. Scammon, University of Utah
Richard Semenik, University of Utah
Terry Shimp, Kent State University
Allan D. Shocker, University of Pittsburgh
Robert E. Spekman, University of Maryland
George Sproles, University of Houston
Ron Stiff, University of Baltimore
John Swan, University of Arkansas
Nancy Thal, University of Colorado
Valerie Valle, University of Pittsburgh
Donald Vinson, University of Southern California
Marjorie Wall, University of Guelph
Robert Westbrook, Duke University
Malcom White, California State University, Sacramento

